

The Concept of Will in Classical German Philosophy

Between Ethics, Politics, and Metaphysics

Edited by
Manja Kisner and Jörg Noller

DE GRUYTER

ISBN 978-3-11-064912-3

e-ISBN (PDF) 978-3-11-065463-9

e-ISBN (EPUB) 978-3-11-065174-4

Library of Congress Control Number: 2020934031

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Angaben sind im Internet über <http://dnb.dnb.de> abrufbar.

© 2020 Walter de Gruyter GmbH, Berlin/Boston

Satz: 3W+P GmbH, Rimpf

Druck und Bindung: CPI books GmbH, Leck

www.degruyter.com

Table of Contents

Manja Kisner and Jörg Noller

Introduction — 1

Part I Kant's Conception of Will

Sorin Baiasu

Free Will and Determinism: A Solution to the Kantian Paradox — 7

Markus Kohl

Spontaneity and Contingency: Kant's Two Models of Rational Self-Determination — 29

Halla Kim

How is the Corruption of the Will Possible? Kant on Natural Dialectic and Radical Evil — 49

Günter Zöller

Eleutheronomy: Will, Law and Liberty in Kant's Esoterically Political Philosophy — 71

Part II The Concept of Will after Kant

John Walsh

The Fact of Freedom: Reinhold's Theory of Free Will Reconsidered — 89

Amit Kravitz

On the Real Possibility of a Pure Moral Will: Maimon vs. Kant — 105

Tom Giesbers

Drive as a Constitutive Element of Practical Action in Jacobi and Fichte — 125

Manja Kisner

Drive and Will in Fichte's *System of Ethics* — 139

Ansgar Lyssy

Reality as Resistance: The Concept of the Will in Bouterwek's *Idea of an Apodictic* (1799) — 159

Jörg Noller

"Will is Primal Being": Schelling's Critical Voluntarism — 181

Daniel Wenz

Hegel's Logical Foundation of the Will: Reconciling Psychology and Social-Ontology — 203

Alex Englander

Hegel and the Paradox of *Willkür* — 227

Jenny Bunker

Ethics and Will in Schopenhauer's Philosophy — 247

Index of Names — 263

Index of Subjects — 265

Notes on Contributors — 269

Manja Kisner and Jörg Noller

Introduction

The concept of will is fundamental to philosophy. It occupies – at least since Augustine – a special place in the history of philosophy. The concept of will plays an important role not only in ethics, but also in moral psychology, political philosophy, philosophy of law, and even in anthropology, epistemology, and metaphysics. However, the concept of will became especially important in the period of Classical German Philosophy. Here, the will evolved into one of the central notions of Kant's practical philosophy, and after that underwent decisive transformations in post-Kantian philosophy. In this short, but prolific time span of about thirty years, each major thinker from Kant to Hegel and Schopenhauer developed his own understanding of the will.

This volume collects thirteen original essays that portray different accounts of the will in Classical German Philosophy and outline its transformations from a historical as well as systematic perspective. Besides articles that shed light on major figures such as Kant, Fichte, Schelling, Hegel or Schopenhauer, lesser-known philosophers such as Bouterwek, Jacobi, Maimon and Reinhold, with their own original contributions to the theory of the will are presented as well. The volume emerged from a conference that took place on November 24th and 25th, 2017, at the University of Munich. Its aim was to bring together both established scholars as well as early career scholars in order to discuss the conceptions of will in and after Kant. Although the presented selection is not complete and many other philosophers could be included as well, the collection nevertheless offers a representative overview of the development of the concept of will in the period of Classical German Philosophy for the first time.

From a systematic point of view, the aim of this volume is to look into this development and to inquire into the relation between the will and related concepts such as “autonomy”, “practical reason”, “drive”, “incentive”, “lower and higher appetitive faculty”, “decision”, “choice”, “intelligible deed”, and “action”. The papers of this volume address and discuss the following questions: How is the concept of will related to the concepts of (practical) reason and understanding? How is the concept of will related to the concepts of drive and striving? How does the concept of will contribute to a theory of (practical) subjectivity? How is the will determined in order to be free? How is the will motivated to morally good action? And finally: How is the (free) will situated within society, history, and the world?

The first part of the volume is dedicated to Kant's practical philosophy and to his conception of the will. In his *Groundwork of the Metaphysics of Morals*

(1785) and *Critique of Practical Reason* (1788), Kant links the will to pure practical reason and presents his notion of autonomy. In contrast to his theoretical philosophy, which focuses on the world of appearances, his practical philosophy deals with the realm of the intelligible, or noumenal. Despite the great importance that the *Groundwork* and the second *Critique* have in his oeuvre, Kant's conception of the will in these two works is not yet unequivocally defined. Kant argues here that "a free will and a will under moral laws are one and the same" (4:447). It is not until *Religion within the Limits of Reason Alone* (1793) and his late *Metaphysics of Morals* (1797) that Kant explicitly defines and points out the difference between *Wille* (will) and *Willkür* (choice). However, the following questions arise in connection with this distinction: Is *Willkür* (choice) just an abstraction of freedom or even a false kind of freedom? Or does our individual freedom involve choosing between the moral law and our empirical motives? What are the motives of our free will? Can they be purely rational or need they be partly empirical? Although Kant's theory of the will is regularly discussed in Kant scholarship, there are still many aspects of his theory that need further elucidation. The main problems concern the solution of the Third Antinomy, Kant's distinction between *Wille* and *Willkür*, the transition from transcendental to practical freedom, and the transition from practical to political philosophy. The four articles on Kant in the first part of the volume approach these problems from different angles by taking into account different works of Kant.

The questions that arise on the basis of Kant's conception of the will were also decisively influential to the subsequent development of post-Kantian philosophy. In order to outline the essential transformations that the concept of will underwent from Kant to Hegel, the following chapters of the second part of the volume discuss the major figures of German Idealism – Fichte, Schelling and Hegel – as well as other more or less well-known post-Kantian philosophers who also developed their own theories of the will under the influence of Kant. One of the early popularizers of Kant's philosophy in general and Kant's practical philosophy in particular was Karl Leonhard Reinhold. His own conception of will, which he developed in confrontation with Kant's theory as presented in the *Groundwork* and second *Critique*, is especially important. Reinhold values Kant's practical account and understands himself as a follower of Kant, but at the same time he also detects some shortcomings in Kant's theory of the will. Accordingly, in the *Letters on the Kantian Philosophy, Volume II*, Reinhold presents his own theory of the will as a critical adjustment to Kant's account. Kant's renewed discussion of the concept of will in his late *Metaphysics of Morals* is – *inter alia* – also due to Reinhold's critique. Reinhold argues for a conception of individual freedom as *Willkür* (choice). This raises the question of how the will is connected to practical reason and how we come to act morally.

Another important, but less well-known critique of Kant's conception of will was developed in the last years of the eighteenth century by Salomon Maimon. Maimon doubts that respect for the moral law is indeed a sufficient driving-force of our finite will. Instead, he argues that moral driving-forces must be rooted in the observable natural driving-forces. On this point, Maimon's account differs significantly not only from Reinhold but also from German Idealists. This discussion about driving-forces (*Triebfedern*) was extremely important for the development of Fichte's practical philosophy. In the period of his *Wissenschaftslehre nova methodo* (1796–1798), Fichte goes a step further than Reinhold and Maimon by not only focusing on the problem of driving-forces, but also by introducing into his system a concept of drive (*Trieb*), which was originally used as a biological term. Furthermore, the concept of drive is important in explaining the concept of will in Fichte's *System of Ethics* from 1798. This book is Fichte's only work on ethics that was published during his lifetime. The concept of drive serves here as an important intermediary concept that connects nature and freedom.

At the end of the eighteenth century yet another, but very little-known philosopher – Friedrich Bouterwek – offered a rival account of the will. On the one hand he followed Kant and understood himself as a Kantian, but on the other hand he advocated a different approach than Reinhold, Fichte or Jacobi. Due to Bouterwek's specific type of realism, he conceives of the will as a natural force that functions in a way similar to other natural forces. With his account he anticipates later, more ontological conceptualizations of the will, such as Schopenhauer's, which, at least to some extent, he might have influenced. But already before Schopenhauer's metaphysics of the will, the concept of will experienced an important shift in Schelling's philosophy towards a more holistic theory of the will that tries to combine realistic and idealistic elements. In a famous passage Schelling identifies the concept of will as primal being (*Ursein*), and thereby points out that the will cannot be properly described in a merely formal way and only with regard to negative freedom, but that it also needs its counterpart in positive freedom. This step toward the identification of will with primal being might be understood as a step in another direction than we observe in Kant, Fichte or Reinhold.

Departing from a merely formal conception of the will, we are also led to Hegel's account of the will discussed in his *Science of Logic* as well as in the *Elements of the Philosophy of Right*. Hegel focuses mainly on the distinction between *Wille* (will) and *Willkür* (choice). Unlike Reinhold, who defended a conception of freedom as *Willkür*, Hegel argues that the concept of *Willkür* fails to unify the form of voluntary choice with its content, which is always given from outside, and is therefore heteronomous and contingent. For that rea-

son, *Willkür* can never truly be a form of rational autonomy. Also, *Willkür* cannot exist on its own, since it needs commitments to ethical life as a background, which Hegel accounts for by means of his conception of *Sittlichkeit*.

Finally, Schopenhauer's philosophy of the will combines different aspects of previous conceptions and transforms them into a metaphysics of the will. According to Schopenhauer, the will is not just a human capacity, but rather a universal principle – a blindly striving, non-rational force that all living beings share. Unlike the previous philosophers, however, Schopenhauer qualifies the striving of the will in negative terms. Similar to Schopenhauer, Schelling and Hegel also point out the dynamic character of the will and conceive it as a striving force, but they differ substantially in how they define the purpose of this striving. Whereas for Schelling, Hegel or Fichte, the main goal of the will still lies in the realization of freedom and morality, Schopenhauer's principle of the will does not have any rational goal and is also never fulfilled. This brings Schopenhauer to his ethical dismissal of the will and to his views on the denial of the will. Schopenhauer's credo is that we need to negate the will in order to escape its meaninglessness.

This brief overview of different conceptualizations of the will in Classical German Philosophy demonstrates how important and much discussed this topic was in this time span. Starting from a merely ethical perspective, the concept of will went through various transformations and gradually evolved into a concept that supersedes mere moral philosophy and develops into an ontology and sociology of the will. The aim of this volume is to present this genealogy of the will by pointing to the vivid and very prolific exchange of ideas between philosophers active in the period of Classical German Philosophy.

The editors of the volume would like to thank Philip Zogelmann M.A. for his editorial support.

Sorin Baiasu

Free Will and Determinism: A Solution to the Kantian Paradox¹

Abstract: It is generally acknowledged that Kant's solution to the problem of free will is beset by various problems. On some accounts, these problems make Kant's solution incoherent, ambiguous and confusing. In this paper, I examine several issues that allegedly threaten Kant's solution and I identify the most significant one, purportedly leading to the incoherence accusation. I then show that Kant's philosophy has the conceptual resources to explain away the inconsistency and solve the so-called Paradox of Kantian Libertarianism.

1 Introduction

I take it to be relatively uncontroversial to assert that Kant is committed to phenomenal determinism. Phenomenal determinism is the claim that all events in space and time are subject to deterministic causation. They are subject to deterministic causation, since all events in space and time fall under laws of nature (that is, behave in accordance with laws of nature) and, hence, they follow, in accordance with those laws, from previous states of the universe. We find such an argument in the Second Analogy.

In the Antinomies, Kant also holds that an alternative causality is possible. This is the causality of freedom. The causality of freedom involves a capacity to initiate a new chain of causes and effects, which is not a determined function of the previous state of the universe under specific laws of nature. More generally, from the Dialectic, we know that we cannot have theoretical cognition of the causality of freedom, since freedom is a transcendent, metaphysical idea. Kant's aim is to show the possibility, and not the actuality, of the causality of freedom. Assuming Kant is successful, his claim is that it is in principle possible that some of

¹ Acknowledgements: An early draft of this paper was part of a keynote presentation I gave to "The Concept of the Will in Classical German Philosophy" conference, held at the University of Munich, in November 2017. I am grateful to the participants to the event for great discussions. I am also grateful to the organisers of the event (and co-editors of the forthcoming collection) for some excellent questions. Part of the work for this paper was carried out under the aegis of the British Academy Newton Advanced Fellowship "Dealing Ethically with Conflicts between Deep Commitments", for which I act as co-holder. I am grateful to the British Academy for their support.

our actions (as events in space and time) be due to the causality of freedom. Call this Kantian libertarianism.

The problem arises, then, that phenomenal determinism and Kantian libertarianism seem to yield a contradiction. On the one hand, my action, as an event in space and time, is causally determined (phenomenal determinism); on the other hand, however, my action is at least in principle free from causal determinism (Kantian libertarianism). Any position in the metaphysics of free will, which rejects compatibilism in search for a stronger notion of freedom is likely to be faced with this contradiction, assuming that phenomenal determinism is also in place. Call this the Paradox of Kantian Libertarianism.

The aim of this paper is to show that Kant has an answer to this paradox. This answer is quite clearly present in Kant's texts, but it is not presented as such by Kant. Moreover, as far as I am aware, it is a solution which has not been presented so far in the secondary literature.²

2 The Imputability of Voluntary Actions

Consider the following claim in the *Critique of Pure Reason*:

Hence reason is the permanent condition of all the voluntary [*willkürlich*] actions under which the human being appears. Each of these actions, even before it occurs, is predetermined in the human being's empirical character. But in regard to the intelligible character, of which the empirical character is only the empirical schema, no before or after holds, and every action – regardless of its time relation to other appearances – is the direct effect of the intelligible character of pure reason. (A553/B581)^{3,4}

2 For instance, Xie (2009) criticises Kant for providing an incoherent, ambiguous and confusing account. His argument is developed by critical reference to Allen Wood (1984) and Henry Allison (mainly 1990, but also 1993). I will present his account in more detail later in this paper. Allais suggests a compromise solution, which relies on the view that science is in principle incomplete (Allais 2015, 305f.). In other words, Kant's determinism is not so strong as to eliminate freedom. This is certainly an ingenious solution, but goes against some of Kant's claims concerning the strict necessity of phenomenal determinism. Although it is not my aim to show this here, I think the solution I will present here avoids the need for a compromise, which weakens Kant's determinism, and can explain also some further problems sometimes identified in Kant (see, for instance, Rohlf 2018, 5.2).

3 In citing Kant's works the following abbreviations are used:

MS: *The Metaphysics of Morals* (*Die Metaphysik der Sitten*), comprising the *Metaphysical First Principles of the Doctrine of Right* (*Metaphysische Anfangsgründe der Rechtslehre*) (1797) and the *Metaphysical First Principles of the Doctrine of Virtue* (*Metaphysische Anfangsgründe der Tugendlehre*) (1797), in Kant (1996: 353–603).

Kant talks about voluntary actions as having reason as a permanent condition. He also says that every voluntary action is pre-predetermined empirically even before it occurs, but intelligibly it is the direct effect of the character of pure reason and it is not determined in accordance with the causality of nature. Kant makes it then clear that by voluntary action he means imputable action and that he takes imputable actions to be imputable in virtue of the fact that pure reason can be an effective cause of voluntary actions (A553/B581ff).⁵

GMS: Groundwork of the Metaphysics of Morals (Grundlegung zur Metaphysik der Sitten) (1785), in Kant (1996, 41–108).

URM: *On a Supposed Right to Lie from Philanthropy* (*Über ein vermeintes Recht aus Menschenliebe zu lügen*) (1797), in Kant (1996, 611–15). Pagination references in the text and footnotes are to the volume and page number in the German edition of Kant's works, *Kants gesammelte Schriften* (1900–) with the exception of references to the *Critique of Pure Reason*, which will be referred to following the standard A (first edition) /B (second edition) convention. I am using the translations listed in the References.

4 Kant talks in this quotation about reason, but it becomes clear that he has in mind pure reason. This will exclude an instrumental concept of reason based on hypothetical imperatives, since these depend on a conditional imperative; by contrast, Kant has in mind in this context pure reason, as “unconditioned condition of any voluntary action” (A554/B582). In what follows, I will use ‘reason’ and ‘pure reason’ interchangeably.

5 “[...] let us take a voluntary action, e.g., a malicious lie, by means of which a person has brought a certain amount of confusion into society. And suppose that we first investigate his action as to its motivating causes from which it arose, and that thereupon we judge how the action can, along with its consequences, be imputed to him. In pursuing the first aim we search through the agent's empirical character until we come to its sources. We locate these in bad upbringing, evil company, partly also in the wickedness of a natural makeup that is incentive to shame; and partly we put them to frivolity and rashness. Here, then, we do not ignore the occasioning causes that prompted the action. [...] But although we believe the action to be determined by these causes, we nevertheless blame the perpetrator. We blame him not because of his unfortunate natural makeup, nor because of his previous way of life. [...] This blame is based on a law of reason; and reason is regarded in this blaming as a cause that, regardless of all the mentioned empirical conditions, could and ought to have determined the person's conduct differently. And the causality of freedom is by no means regarded merely as concurrence; rather, it is regarded as in itself complete, even if the sensible incentives were not all for this causality but were even against it. The action is imputed to the agent's intelligible character: now – at the instant when he is lying – the guilt is entirely his. Hence his reason, regardless of all the empirical conditions of the deed, was wholly free, and to its failure is the deed to be imputed entirely. [...] Reason is present to, and is the same in, all actions of the human being in all circumstances of time. But reason itself is not in time, and by no means gets into a new state in which it previously was not; with regard to this state reason is *determinative*, but not *determinable*. Hence we cannot ask, Why did reason not determine *itself* differently? – but only, Why did reason not determine *appearances* differently through its causality? To this, however, no answer is possible” (A555 – 6/B583 – 4).

One clarification is in order here. When Kant says that every voluntary action is the direct effect of the intelligible character of pure reason, he includes actions which, as well will see, pure reason fails to influence. So this direct effect that the intelligible character has on all actions should not be understood as the effect of pure reason. Kant only makes reference here to the fact that all voluntary actions are actions which pure reason *can* completely determine and, hence, actions will be different depending on whether pure reason acts as a complete cause or fails to act. As the intelligible character refers to the laws which govern pure reason and as all voluntary actions are, through pure reason, under the direct governance of these laws, they are also the direct effect of the intelligible character.

Consider Kant's illustration, which follows the above quotation: a malicious lie by which a person has successfully deceived some others. Kant imagines first an investigation into the motivating causes of the lie: bad upbringing, evil company, the wickedness of a natural makeup insensitive to shame, frivolity and rashness. He thinks that, although we do believe the action to be determined by these causes, we nevertheless blame the perpetrator. The reason for the blame, Kant notes, does not have to do with this bad upbringing, evil company, wickedness of natural makeup, frivolity or rashness, since we presuppose we can set aside "entirely" his previous way of life. Instead, we regard this person as able to produce a series of consequences without preceding motivating causes, completely on his own.

Thus, Kant notes, the blame is based on a law of reason. Reason is regarded as a cause which, regardless of all the motivating conditions, could, and ought to, have determined the person's conduct differently. Moreover, Kant adds, reason is not a concurrent cause – in other words, it is not simply the case that the motivating causes and reason exerted together a force in the direction of the considered effect (the malicious lie). Reason is regarded in itself as a complete cause: if the motivating causes go against the causality of reason (as in this case), reason could and ought to still determine the person's conduct. Moreover, Kant adds, the deed (the malicious lie) is imputed entirely to reason's *failure* to do so.⁶

⁶ On the basis of Kant's text in the first *Critique*, it is I think uncertain whether Kant thinks that reason always exerts a causing influence on our actions or whether he thinks that reason does so only for certain actions (e.g., a truthful statement, when motivating causes encourage the agent to state a malicious lie), but not for all (e.g., the malicious lie, where reason failed to intervene). The actions can still be imputable anyway: in the first case, on account of freedom, in the positive sense – the action was determined by reason, which determined the agent to refrain from uttering a malicious lie or to make a truthful statement, depending on the manner in which

The model of imputability Kant seems to use here is plausible: we may be accountable for what happens even in cases where what happens is not the direct result of our actions, but it is the result of our passive attitude, in circumstances where we had the possibility to effect change. The fact that I passively regard the course of events without intervening (even when such events are the result of behavioural patterns I have and which I allow to manifest themselves) does not mean that I cannot be accountable for the outcome.

There are two issues which I think it is important to distinguish for the purpose of this paper. First, there is the issue of holding somebody accountable for her behaviour. Kant's answer to this, as we have seen, is that voluntary behaviour is a condition of accountability, where voluntary behaviour is behaviour which the agent could have changed independently from its motivating causes. Presumably somebody who is hypnotised into lying will not be held accountable for her lies, since she could not have acted differently under the spell of hypnosis.⁷ The second issue concerns the direction of responsibility. Kant thinks that a law of reason guides us into blaming or praising a person for her behaviour (assuming it is voluntary behaviour). The person who does something she ought not have done will be blamed and the blame is imputed to reason's failure to prevent the behaviour (in this case, the malicious lie) from happening.

The reason why it is important to distinguish between these two issues is that for the case of 'actions' for which the person is not accountable, the Paradox of Kantian Libertarianism does not apply.⁸ What a person may be doing without being aware of doing it or when she is compelled to move in a particular way does not raise a problem for Kant's account of freedom, since the person is not free. In what follows, therefore, the focus will be on actions for which a per-

it determined appearances; in the second case, for the malicious lie, where reason supposedly had failed to exert any causal influence, on account of freedom, in the negative sense – the action was not necessitated by the motivating causes, which played a determining role. In this paper, I adopt the second interpretation both on the basis of what Kant says here and on the basis of some of his claims in the writings of practical philosophy; it is not important to establish conclusively the accuracy of the interpretation, since the aim is just to show that there is a solution to the Paradox of Kantian Libertarianism within the framework of Kant's philosophy.

⁷ Assuming, of course, she does not consent to being hypnotised and being suggested to lie.

⁸ The reason why I place the word 'action' in inverted commas is that I think unaccountable 'actions' are not really actions, but movements made by a person, who is neither directly nor indirectly in control of what she is doing. For instance, the person who is hypnotised is not directly in control of her bodily movements. There is of course the classical argument according to which she might be indirectly in control of her movements, through the consent given to being hypnotised. Yet, I assume here that, when a person is neither directly nor indirectly in control of what she is doing, then she cannot be held accountable for those movements.

son is accountable. On Kant's account, it seems that, in principle, a person is accountable for her behaviour, even if her reason has played no role in the determination of her conduct. The condition for accountability is the *possibility* for reason to determine conduct, but not its actual determination. Given that, according to Kant, reason has the capacity to determine conduct independently from other motivating factors, perhaps the best way to explain accountable conduct is as conduct where reason *can* determine action – irrespective of whether eventually it does so entirely or it does not intervene at all (lets the motivating factors determine the action) or does something in-between.

Now, the question of when reason can determine conduct (as in Kant's case of the malicious lie) and when it cannot (as in my example of the hypnotised person and the refined version of the person who was hypnotised without her consent) is an interesting one, but it is not one I will focus on in this paper. I assume (without failing to see all the difficulties which need to be addressed, if such an assumption were to be defended) that we have a way of determining this fairly reliably. The question now concerns the nature of the difficulty introduced by the Paradox of Kantian Libertarianism in Kant's discussion of accountable conduct. More exactly, it may seem that Kant contradicts himself when he says that “we believe the action to be determined by these [motivating] causes” (A555/B583), on the one hand, and, on the other, that “we nevertheless blame the perpetrator” (A557/B585). I will examine this difficulty more closely in the next section.

3 Some Difficulties of Kantian Libertarianism

To see where the source of the contradiction might be, consider the following account, which concludes with these claims about Kant's account of freedom: “[Kant's] solution to the free will problem [is] incoherent, ambiguous and confusing [...]. [H]is solution can be regarded in general as a failure” (Xie: 2009, 76).⁹ This is Xie's account of Kant's solution to the problem of Kantian libertarianism – call this, the Anti-Kantian Account.

⁹ In fairness, these claims are qualified: Kant's solution “[...] can also be regarded as a success ‘in the context of his philosophy’ [...]”. Kant may not, however, take all the blame for his failure because a successful solution to the free will problem may not have been available in his time” (Xie 2009, 76). To be candid, however, I am not sure how these qualifications speak in favour of Kant – first, it is unclear how valuable an account is, when it is successful in the context of the author's philosophy, but it is “ambiguous”, “incoherent” and “confusing” in general; secondly, it is unclear to me how the author of an ambiguous, incoherent and confusing answer to a prob-

First, Xie notes that ‘compatible’ refers to the natural agreement of two things (whether concepts or persons). By contrast, two things are *incompatible* when they disagree with each other naturally and cannot coexist harmoniously naturally, although they may be made compatible under certain specific conditions. Yet, freedom and determinism are incompatible terms, if determinism is taken to mean determination by antecedent causes, whereas freedom, non-determination by antecedent causes. One aspect which I would like to mention, since it is not explicitly noted as part of the Anti-Kantian Account, is that, when the metaphysically compatibilist think freedom and determinism are compatible, they regard it in this way, because they do not consider freedom as such a non-determination, but as determination by a particular antecedent cause, namely, sensible desire. Hence, compatibilism does not provide a solution to the problem of Kantian libertarianism, but to the different ‘issue’ of how determinism is compatible with actions based on desires, which is not in fact an issue at all (that is, not an issue for Kant’s account).

On the Anti-Kantian Account considered here, the notion of determinism usually employed in the freedom versus determinism debate is the empiricist notion, according to which the link between cause and effect is not a necessary one, but is a customary belief we usually have about the two events. As such, the claims which are usually taken to be made in the metaphysical debate are the following: compatibilism is understood as making the claim that, if determinism is true, then libertarianism will also be true, whereas incompatibilism is understood as the claim that, if determinism is true, then libertarianism will be false. By contrast, from the rationalist perspective, we can understand compatibilism to be the claim that, if libertarianism is true, then determinism will also be true and, more significantly, we can understand incompatibilism to make the claim that if libertarianism is true, then determinism will be false.¹⁰

lem is to be blamed only if a successful solution were available – could the author not be blamed for offering an account which is not at least precise, coherent and clear, even if not also a successful solution to the problem under consideration? Finally, I wonder what Xie’s opinion of Kant (or perhaps more generally of philosophers) is, if he is only expecting him (them?) to offer solutions to particular problems, if the problems already have solutions.

10 Interestingly, Xie does not seem to notice (Xie 2009, 67) that, if we take the empiricist notion of determination, then ‘freedom’ and ‘determinism’ are not incompatible, since I can be free (in the sense that my action is without antecedent causes), while at the same time being determined (that is, being *believed* to perform actions with antecedent causes). Of course, Kant’s account of causation is different – at phenomenal level, causes and effects are necessarily connected and not merely as a matter of customary association. So, what Xie should probably have said is not that the starting point in the free will debate is an empiricist notion of determinism, but that it is simply determinism.

Given the empiricist undertones of the terms ‘in/compatibilism’, the Anti-Kantian Account replaces them with ‘hard/soft determinism’, respectively. Kant can be seen as a phenomenally hard determinist, but noumenally as a libertarian. Kant makes freedom the primary causality and nature only the secondary causality grounded in freedom. Freedom and determinism do not become compatible, since they are still contradictory, incompatible terms. Kant makes them compatible by ascribing them to different worlds, and by subordinating nature to freedom; moreover, he conceives of the effects of freedom as also determined by natural causality.¹¹ According to Kant, a subject belonging to the sensible world has an empirical character; the same subject, as belonging to the noumenal realm, has an intelligible character.

The problem, then, is how to tell when the subject belongs to the sensible world and when, to the intelligible. This is a problem, because:

Surely the same subject cannot belong to the two worlds *at once*, if it can, then the idea of belonging to two worlds will become meaningless because in that case, we can also say that the same subject belongs to three, five or a hundred worlds at the same time since we do not have to tell on a particular occasion that the subject belongs to *this* world rather than *that* world or has *this* character rather than *that* character. It is in explaining this point that Kant may have confused his readers more than at any other points. [...] For whatever reason, Kant’s failure to explain that point clearly is inexcusable. (Xie 2009, 72–3)

The claim here is that the same subject cannot belong to two worlds at once. The argument seems to be that, if the same subject belonged to two worlds at once, there would be no need to specify on particular occasions to which world she belonged and, hence, she could as well be taken to belong to three, five or a hundred worlds. At this point, I do not want to question any of these, but only to present the argument and clarify any presuppositions or implications. As I will briefly mention later, in my view this problem disappears once we take a Two-Aspect, as opposed to a Two-World interpretation of Kant’s distinction between phenomena and noumena.

Next, on the Anti-Kantian Account considered here, the problem which Kant seems to have, as far as the Paradox of Kantian Libertarianism is concerned, emerges from his claims in the first quotation given above, at the beginning of

¹¹ It is unclear why Xie thinks Kant needs to make freedom primary and why he needs to make the effects of freedom determined by natural causality. He claims these would be needed to make freedom and nature co-exist, but I think they are able to co-exist once they are made to belong to different worlds. I should add that I am sceptical of the use of ‘world’ in this context – I think it should be made clear that Kant does not adopt a Two-World account of the distinction between phenomena and noumena, but a Two-Aspect account.

Section 2. There, Kant suggests that the intelligible character has effects which belong to appearance. He also maintains that appearance is determined by the empirical character. In addition, he takes the empirical character to be the appearance of the intelligible one. Furthermore, he sees certain actions as only caused by the empirical character (although the empirical character itself is determined by the intelligible character). Yet, the empirical character can also cause its own series of events in the time sequence; the intelligible character can cause effects in the phenomenal world independently from the empirical character, effects which are merged into the time sequence of appearances.

What these claims suggest is that we could know which effects are caused by the intelligible character and which by the sensible character, if we viewed the phenomenal world from the intelligible perspective. Yet, for the Anti-Kantian account, Kant seems to have two problems which make any such identification impossible. First, he allows us to have intelligible character, but he thinks we cannot know this character. This point seems confirmed by Kant's claim that

[w]e could not, indeed, ever become acquainted with this intelligible character directly, because we cannot perceive anything except insofar as it appears; but we would still have to *think* it in accordance with the empirical character, just as in general we must – in thought – lay a transcendental object at the basis of the appearances although we know nothing about this object as to what it is in itself. (A540/B568)

The use by Kant of 'directly' or 'immediately' [*unmittelbar*] can be questioned here, since it implies that we can know or become acquainted with the intelligible character indirectly.¹² We can also question the sense of his claim that we know nothing about the transcendental object as to what it is in itself – as we can know that the intelligible character is good or bad, why is this not sufficient for our knowledge of the intelligible character? While Kant might well be right that we cannot know why the intelligible character is good or bad, the fact that we know that it is in one way or the other should be sufficient for us to say that we do know the intelligible character.¹³

For the Anti-Kantian Account, there is a second obstacle to an identification of the effects produced by the intelligible and sensible character. This is an over-

¹² I use both 'acquainted' and 'known', because Kant's first *Critique* has, in the Academy edition, 'gekannt', where, according to Erdmann, should be 'erkannt' – see n1 on p. 528 of the 1956 Felix Meiner edition.

¹³ It is unclear to me whether there is any problem here: Kant acknowledges that we become acquainted with the intelligible character indirectly. This should enable us to offer a supported judgement about whether the character is good or evil. This enables us to claim that the character is good or evil, but not to know how the character is in itself.

emphasis on the empirical character. Kant takes the empirical character to be the empirical cause of all our actions; if so, how can the causality of the intelligible character play a role? As we have seen, on the one hand, Kant claims that “reason is the permanent condition of all the voluntary actions under which the human being appears” and “[e]ach of these actions, even before it occurs, is predetermined in the human being’s empirical character” (A553/B581); on the other hand, however, he also thinks that “pure reason acts freely, i.e., without being dynamically determined in the chain of natural causes by external or internal bases that precede the action as regards time” (A553/B581). If a person acts always under the condition of freedom, then does she not act freely and, therefore, under the causality of her intelligible character? If so, however, how can her actions be predetermined in her empirical character too?

Although the empirical character is the appearance of the intelligible character, this does not explain how a person’s actions can be free and determined at the same time. If reason is free, then its effects should belong to the intelligible character; if reason is unfree, then its effects should belong to the empirical character. By grounding empirical character in the intelligible character, Kant (it is claimed as part of the Anti-Kantian Account examined so far) should not confuse the effects coming from these two distinct sources. The empirical character has become a separate identity, so it must only determine what directly comes from itself. This will allow the effects coming from the intelligible character to be determined by the intelligible character. This leads to the conclusion that Kant’s account is contradictory and inconsistent:

[t]his is what makes Kant contradictory and inconsistent regarding the relationship of appearances to the intelligible and empirical character, and hence makes his view confusing to his readers. (Xie 2009: 75)

On the Anti-Kantian Account, the origin of the problem is the relation between the empirical and intelligible character. The possible reply that makes reference to Kant’s agnosticism is not considered very strong. According to this reply, Kant has to admit that free action out of reason must be the effect of the intelligible character; yet, given his agnosticism, we cannot know the intelligible character from its effects, so we must also be able to see these effects as determined by the empirical character, which is all we, human being, can know. The reason why this reply is not very strong is that, according to the Anti-Kantian Account, when an agent acts freely, the outcome (the free actions performed by the agent) is the effect of the intelligible character. When the same agent acts according to the natural law, there will be the effects of her empirical character only. Yet, whereas Kant does separate causality, he does not separate the effects.

On the Anti-Kantian reading considered so far, Kant himself seems to acknowledge that it is impossible for us to distinguish between these effects:

Hence the morality proper of actions (merit and guilt), even the morality of our own conduct, remains entirely hidden to us. Our imputations can be referred only to the empirical character. But no one can fathom how much of this character is a pure effect of freedom, and how much is to be ascribed to mere nature: viz., either to a defect of temperament that one has through no fault of one's own, or to one's temperament's fortunate constitution (*meritum fortune*). And hence no one can pass judgement in accordance with complete justice. (A551/B579 n)

Here, Kant confirms that the extent to which an action is caused as part of the empirical or intelligible character is not precisely ascertainable by us. Hence, a morally right action may be the effect of a fortunate constitution, which makes the agent's merit comparatively smaller than that of an agent who must withstand a defect of temperament in order to perform the same action. Yet, according to the Anti-Kantian Account, in principle, it should be possible to distinguish between the following: (a) effects caused by the empirical character only; (b) effects caused by the intelligible character only; (c) effects caused by both characters, predominantly by the empirical character; and (d) effects caused by both characters, but predominantly by the intelligible one.

I think, however, that this is not the main problem of Kant's account, for, assuming we would be able to accurately identify and distinguish between (a), (b), (c) and (d), the Paradox of Kantian Libertarianism would be left untouched: if an action is caused entirely or predominantly by the intelligible character, then it is free or to a large extent free; yet, on Kant's account, the same action is also (entirely) predetermined by the empirical character and, hence, not free. Here the problem is simply one of what looks like contradictory claims made by Kant; it is not a problem concerning our ability to ascertain which character caused a particular action and to what extent – these are judgements of relative merit¹⁴.¹⁵ Assuming that I accurately identify, say, the action entirely caused

¹⁴ In fact, I think it would be better to call them judgements of relative desert or deservingness. Merit is usually distinguished from desert as a notion which refers to entitlement, that is, to the extent to which an agent meets the conditions for a particular reward, irrespective of whether that reward is deserved. For instance, the shopkeeper who happened to give the right change to her customers over one month may be entitled or merit the prize of best shopkeeper of the month, but may not deserve it, if always giving the right change is the result of an error she made in her calculations. See, for instance, Baiasu (2006, 229f.)

¹⁵ And, yet, this is what Xie claims: "As to the question I posed earlier, which effects are caused by which character, it is a very difficult question however, it may not be impossible to answer, as

by the intelligible character, we will still have to face the question of how this action can at the same time be predetermined by the empirical character and follow necessarily from some natural causes under specific conditions. To answer this (and hence, the Paradox of Kantian Libertarianism) will be the focus of the next section.

4 My Solution to the Paradox of Kantian Libertarianism

As we have seen, according to Kant, we can talk about blame and a perpetrator, when an action is imputable, and it is imputable, when it *can* be completely determined by reason. Kant contrasts actions which are determined by reason with actions determined by motivating causes (for instance, bad upbringing, evil company, wickedness of natural makeup, frivolity or rashness¹⁶). Yet, both actions which are determined by reason and actions which are determined by motivating causes are imputable on Kant's account. One question, therefore, is the following: if action A is determined by motivating causes, then A follows necessarily from those motivating causes; at the same time, A is imputable and, hence, possibly entirely determined by reason – hence, how could reason have determined the agent to act differently, if the action is already determined by motivating causes? If A were determined by motivating causes, would it not follow that the agent could not have acted differently and, hence, that reason could not have completely determined the agent not to perform A?

The key claim, which suggests an answer to this problem and which I have mentioned above, is Kant's assertion that, in the case of the malicious lie, which was determined by motivating causes, reason failed to exercise its causal power. As we have seen, imputable or accountable actions are those which *can* be determined by reason independently from motivating causes. When reason lets motivating causes determine the action (this being the case illustrated by Kant with the example of the malicious lie, where reason fails to prevent the malicious lie), the action still counts as voluntary, although there is no contradiction between

Kant humbly claims. The convincing answer to this question may prove vital for solving all the other controversial issues involved in the free will problem" (Xie 2009, 76). As suggested above, I think this misidentifies the problem raised by the Paradox of Kantian Libertarianism.

¹⁶ Of course, our action can also be determined by good upbringing or good company or fortunate natural makeup, etc. The negative motivating reasons he offers are those possibly linked to the case in his example, that of a malicious lie.

the fact that we take the motivating causes to determine the action and the fact that the action is voluntary: although reason *can* determine the action independently of those motivating causes, because it does not intervene, those motivating causes are the determining grounds of action.

There is a potential issue here: if the action is indeed determined by motivating causes, then it must follow necessarily from them under the relevant circumstances, in which case it is unclear how reason could have prevented this action from taking place or could have determined another action. The answer, however, is in Kant's Third Antinomy. There, Kant shows that the hypothesis of the existence of only natural causality in our account of what happens in the world leads to contradictions; he also shows that the contradiction between the thesis and the antithesis can be reconciled by accepting that it is possible for the agent to act freely, that is, independently from natural necessity. Within the context of phenomena, the only factor accounting for what happens in the world is natural necessity. This type of necessity, however, as given by causal laws, is dependent on the *a priori* conditions of our experience. The causality of the necessity of freedom, by contrast, is not conditional in this way. For, as specified by Kant, we get beyond the context of phenomena, when we consider things as they are in themselves, and not simply as we are experiencing them. What may seem a necessity from within the more limited perspective of phenomena may not be from the enlarged context of noumena.

This, by the way, shows that the discussion of the Paradox of Kantian Libertarianism does not need a Two-World interpretation of Kant, as assumed by the critical account presented in the previous section. There is no need to assume that there is a phenomena world and a noumenal world and that there are some corresponding relations between the entities in the two worlds, such that for every phenomenon there is a thing in itself in the noumenal world which is somehow connected with that phenomenon.¹⁷

Going back to distinction between the necessity of the laws of nature and that of the laws of freedom, we can think of an analogy along the following lines: consider a person (call her the uni-person), whose sensibility is structured by a unidimensional *a priori* intuition of space (see Fig. 1);

¹⁷ Once this is acknowledged, the problem of how to identify which world we are in, a problem mentioned above as raised by the Anti-Kantian Account, no longer makes sense.

Unidimensional
Spatial Intuition

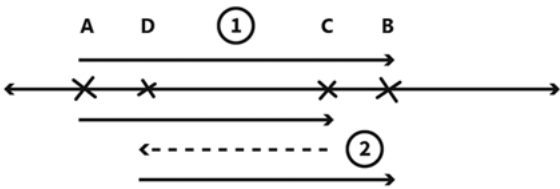


Fig. 1

in this case, to go from A to B is an action, which is *necessarily* given by a limited set of options – say, going directly from A to B or going to C, returning to D and the carrying on to B. Yet, in whichever way going from A to B will happen, it will *necessarily* happen along the only dimension available for the uni-person. Consider now, however, that the uni-person learns about Kant’s philosophy and about her a priori intuition of space and thinks of persons who may have a different type of sensibility, say a bi-person whose sensibility is structured by a bidimensional intuition of space (see Fig. 2).

Bidimensional
Spatial Intuition

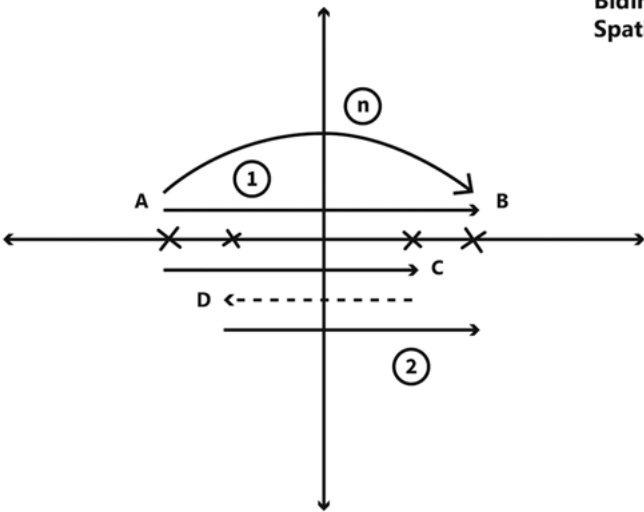


Fig. 2

For this second type of person, possibilities suddenly multiply exponentially, since there are now so many ways in which the journey from A to B can take place. The necessity of the available options for the uni-person is conditioned by the a priori structures which make possible her experience. With bi-persons, we consider the same task, but from a different perspective and the necessities which were in place for uni-persons no longer apply.

Consequently, when we consider the malicious lie as a case of an action where reason failed to exercise its causal power, the context for the determination of the action is that of natural necessity, where the action is determined by motivating causes. That, in that context, the action is determined by natural necessity is not in contradiction with the possible determination of the action by reason in a context in which we consider the action as it is in itself, independently from the way we experience it and, hence, independently from natural causation. When the necessity of reason fails to activate (irrespective of why this is happening), the action is subject to the necessity of nature. In other words, an action's being determined by natural causation in the realm of phenomena is compatible with its possible determination by the causality of freedom, when we take into consideration the necessity of freedom.

The more difficult problem, however, seems to arise not when reason is failing to act and lets natural causality determine the agent's action, but when it acts and brings about an action entirely or predominantly determined by reason. For it is difficult to understand how this action can also be determined by natural causality, given that it is a distinct action compared with that which would have been determined by natural causality. To illustrate this in the context of Kant's example: it seems clear how it is possible for the malicious lie to be seen as determined by natural causation (the necessitating power of motivating causes, such as bad upbringing and evil company), while at the same time being potentially determinable by reason (which in this case fails to exercise its power): the action reason allows to be performed is the action naturally determined by motivating causes and it is an action imputable to the agent, who could have acted differently. But if reason does intervene and causes the agent's silence by preventing the utterance of the malicious lie or perhaps causes the agent to utter part of the truthful statement, then it is unclear how such actions are also caused by motivating causes.

This shows that the main problem captured by the Paradox occurs when we consider actions actually determined by reason.¹⁸ This is because, when reason

¹⁸ To avoid debates on the issue of the grounding of action, I assume here it is at least in principle possible for reason to determine partially an action. In other words, I will consider the case

determines at least part of the action, that part is going to be different from how it would have been had it not been determined by reason. Again, our person who is motivated to utter a malicious lie by bad upbringing, evil company, wickedness of natural makeup, frivolity or rashness would end up silent or truthful, if reason determined her action. For reason, on Kant's account, would justify a prohibition of lying and, hence, would command at least a quiet attitude, if not a truthful one. We seem to end up here with a stark contrast between one action (malicious lie) and another one or a lack thereof (however we would want to interpret keeping silent).

Hence, the problem can be stated as follows: on Kant's account, on the one hand, we have to believe the action (keeping silent, say, in response to a question) to be determined by sensible motivating causes¹⁹ and, yet, on the other hand, we have to accept that it is possible for reason to determine the same person not to lie and, hence, for the same action to be determined by reason. Here, perhaps more than in any other circumstances, the problem of an inconsistency in Kant's account of free will comes to the surface. Yet, I think Kant has available the conceptual means needed to address this worry. To see this, consider the following account of action in accordance with the Categorical Imperative.

I think it is not very controversial to identify two types of action, which is in accordance with the Categorical Imperative. There is first that action which is in accordance with the Categorical Imperative, but is not performed for the sake of the Categorical Imperative; secondly, there is the action which is performed both in accordance with, and for the sake of, the Categorical Imperative. For instance, refraining from uttering a malicious lie or stating part of the truth²⁰ can be done for other motivating reasons (for instance, to avoid the likely consequences of possibly being caught lying or to benefit from some probable positive outcomes of eventually being proven honest). It can also be done because this is the right thing to do and irrespective of other motivating reasons, if we assume here Kant is right when he claims that reason can be a complete cause of action.

where reason has some influence on the action. Since 'some' can mean more than nothing and potentially everything, this will cover a broad range of cases; yet, if reason can partially determine action, then we will have a potential contradiction between Kant's phenomenal determinism and libertarianism only by reference to that part which is determined by reason.

19 As we have seen, according to Kant, "[e]ach of these actions, even before it occurs, is predetermined in the human being's empirical character" (A553/B581).

20 Given Kant's distinction between stating the truth and being truthful (URM 8: 426; for discussion, see Baiasu 2017, 139–41), it would be more accurate to say that the alternative to keeping silent is not to state part of the truth, but to express part of what the agent truthfully thinks is the case, where stating everything amounts to a candid attitude. For the distinction between morality and legality, see, for instance, GMS 4: 407 and MS 6: 225.

Consider now the first case, namely, the case of an action which is performed in accordance with, but not for the sake of, the Categorical Imperative. In this case, it is unclear there is any contradiction, let alone difficulty: the action can still be accounted for as determined by natural necessity, since it is still the result of motivating causes ultimately determining the agent which action to perform. The problem here is not with Kant's account, but with the way I initially described this case as one where reason is the cause of the action. Reason may play a role (for instance, it may command that the malicious lie be avoided), but it is not reason as a complete cause which determines the action. Reason fails to intervene to the end in determining the action, and, for whatever reason, it allows motivating causes to decide the course of action.

Here the important question would be how it is possible for two distinct actions to be regarded as both causally determined by motivating causes under the same circumstances. The answer, however, is that circumstances are not the same: although in both cases reason fails to completely determine the agent to act, in the second case I suggested that reason might still command that the permissible action be one which avoids lying, and this can make certain motivating reasons (namely, those which support acting in accordance with the Categorical Imperative, which Kant does not mention in his account of the malicious lie) visible – there can be likely detrimental consequences of being caught lying, for instance, when lying is seen as wrong. The only question is why in one case reason fails to manifest itself in any way, whereas in the other, it at least issues a command concerning the moral impermissibility of a malicious lie. Yet, this is precisely the question Kant readily acknowledges that we cannot answer, since it concerns determining grounds for reason, that is, determining grounds for a faculty which is supposed to function beyond the phenomenal realm.

The case of the action which is morally worthy (the action performed in accordance with, and for the sake of, the Categorical Imperative) is distinct: here reason does exert its authority both in determining the course of action, which is in accordance with the Categorical Imperative, and in providing the ground for the performance of the action. Here reason is indeed the complete cause of action and it may lead to, say, silence. Yet, in this case, too, the action (e.g., keeping silent) is supposed to be determined by natural necessity and, more exactly, by motivating causes. In the case of the action merely in accordance with the Categorical Imperative, it was quite easy to see how the action would be determined by motivating reasons, since the determining ground of the action was not reason, but precisely such motivating causes, as the desire to avoid the likely detrimental consequences of possibly being caught lying. Yet, in the case currently under consideration, we know that reason acts as a complete cause which determines an action, which is at the same time performed in accordance with, and

for the sake of, the Categorical Imperative. So it seems at least *prima facie* puzzling that the same action is also supposed to be determined by motivating causes.

The first thing to note here is that, if we focus on the action itself and leave aside for the moment the motivation with which it is pursued, we start from the same action both for the case of mere accordance with the Categorical Imperative and for the case where this accordance is accompanied by moral worth. As noted by Kant, the difference between an action which has mere legality and an action which has morality is only the motivation with which they are pursued. So, if we focus on the action of keeping silent, then between the two cases the difference is only given by motivation. If we leave aside the action's motivation for the moment, the actions in the two cases will be identical.

The second thing to note is that the set of potential reasons for action in both cases are also the same. In the case of the action with mere legality we have all the motivating reasons and the motive of duty (since reason commands the action, even if it is not the determining ground of the action). It is, therefore, quite easy to see that, considering the action as a phenomenon, we can provide the same account of the action as determined by motivating causes in both cases (that is, both in the case of the action with mere legality and in that of the action with morality). The only difference is that, in the case of the action with morality, the determining ground is reason, whereas in the case of the action with mere legality, the determining ground is given by motivating causes.

This may seem to be a crucial difference, since we are talking about the determining ground of action. The question is whether, in the case of the action with morality, Kant can claim that the action is also determined by motivating causes or natural causes, when we also know that the determining ground of the action is reason. The worry here is that we end up with another contradiction: on the one hand, the claim is that the action, regarded as a phenomenon, is determined by motivating reasons (on the basis of an account similar to that provided for an action with mere legality); on the other hand, however, we know that the action is in fact caused by reason and that the motivating causes are not sufficient for (indeed, have not actually determined in this case) our morally worthy action.

I think, however, that this worry is easily answered, because, in fact, as we have seen in the discussion related to Kant's Third Antinomy, motivating causes are never sufficient to determine a voluntary action when we take into consideration the broader context of things as they are in themselves. If, in the case of the action with mere legality, we take the account of how the action is causally determined by motivating reasons to be correct, then the same should be done in the case of the morally worthy action. In the case of the action with mere legality,

the agent could have acted differently and this suggests that the motivating causes are not sufficient to determine the action performed by the agent. We need to add also that, in that case, reason failed to exert any influence. Similarly, in the case of the morally worthy action (or the action with morality), again, the motivating reasons are also insufficient to determine the action, but in the phenomenal realm the account is sufficient as a determining cause of the action.

The reason for this peculiar conclusion was already provided: within the limited context of the phenomenal realm, motivating reasons can be seen as necessarily leading to a particular outcome – the malicious lie, for instance; once we move beyond this context, and examine how the action might be determined independently from the particular way in which we experience our agency, we see that the set of motivating causes is not sufficient as a determining cause of action and does not necessarily (this time with a different sense of necessity) have as an outcome that particular action (unless of course, reason fails to exert any influence or only issues a command without attempting to observe it). In short, the crucial aspect of Kant's solution is that provided in the Third Antinomy: we cannot reduce the world to natural causality or to natural causality produced by an uncaused cause; instead, we need to place the uncaused cause in the realm of things in themselves and natural causality in the realm of phenomena.

Once all this is taken into consideration, I submit, Kant's account no longer appears as incoherent, ambiguous and confusing. If this is right, then Kant's philosophy has the conceptual means necessary for solving what I have called the Paradox of Kantian Libertarianism.

5 Conclusion

In this paper, I have examined a set of difficulties usually associated with Kant's solution to the problem of free will. I have presented some of the claims Kant makes, which lead to these difficulties. I have then evaluated the significance of these difficulties and concluded that the main issue seems to be generated by a contradiction in Kant. Very briefly, the contradiction is between the view that the same action is determined by natural causes and that it is free from natural determination. This is the core of what I have called the Paradox of Kantian Libertarianism.

I have, then, identified several cases, which do not seem to raise a problem for Kant's account. First, non-imputable actions are actions where the agency of the person who performs the action does not manifest itself (think of the person who lies under the influence of hypnosis). In this case, there is no contradiction: the action can be explained in terms of natural causality, but there is no claim

that this is at the same time a free action (or even that it is an action, if we understand ‘action’ as referring to imputable agency). Secondly, imputable actions which do not have what Kant calls legality or only have mere legality are also actions where the contradiction of the Paradox does not arise: in both cases, reason fails to act as a complete cause of action, so, as a phenomenon, the action can be accounted for in terms of natural causality.

Next, I have identified the major issue which seems to follow from Kant’s solution to the problem of free will: the case of actions which have both legality and morality. Since this type of action is completely determined by reason, it is unclear how it can also be determined by motivating causes. My solution to this problem relied on Kant’s discussion in the Third Antinomy and some of the claims he makes in his practical philosophy. Thus, the difference between the action with mere legality and that with morality is only given by the ultimate motivating ground, which is duty for the latter kind of action. Duty, as a motivating cause, however, is not part of the motivating causes we identify when we consider the action as a phenomenon; duty is a motivating cause which we consider when we regard action in itself, independently from the way it is experienced by us. This motivating cause, which enables reason to play the role of a complete cause, is no longer conditioned in the way in which the motivating causes of the phenomenal action are conditioned.

Hence, within the context of the action as phenomenon, motivating causes determine the action, but the necessity with which the action is produced holds only within the phenomenal context given by the a priori structures of experience. When we move beyond this context and consider the action as it is in itself, it is possible that the action be actually determined by duty. The necessity of this determination, however, is distinct from that of the motivating causes within the phenomenal context. An account of an action with mere legality in terms of motivating causes is sufficient to explain how the action is produced. Hence, the same account should also be sufficient for the action which is phenomenally the same as the action with mere legality, but which is actually determined by reason. There is no contradiction, therefore, between the fact that, as phenomenal, the action is necessitated by motivating causes, and the fact that, beyond the phenomenal condition, the action is no longer so necessitated, but is the result of reason’s exerting its influence as a complete cause.

Bibliography

Allais, Lucy (2015): *Manifest Reality: Kant’s Idealism and His Realism*. Oxford: Oxford University Press.

- Allison, Henry (1990): *Kant's Theory of Freedom*. Cambridge: Cambridge University Press.
- Allison, Henry (1993) "Kant's Theory of Freedom: A Reply to My Critics." In: *Inquiry* 36(4), pp. 443–464.
- Baiasu, S. (2006): "Debate: The Normative Pluralism of Desert." In: *The Journal of Political Philosophy* 14(2): 226–237.
- Baiasu, S. (2017) "Political Dissimulation à la Kant: Two Limits of the Sincerity Requirement." In: S. Baiasu and S. Loriaux (eds): *Sincerity in Politics and International Relations*. London: Routledge.
- Kant, Immanuel (1900-): *Gesammelte Schriften*. Ed. by the Königlich Preußischen Akademie der Wissenschaften, subsequently Deutsche, now Berlin-Brandenburg Akademie der Wissenschaften (originally under the editorship of Wilhelm Dilthey). Berlin: Georg Reimer, subsequently Walter de Gruyter.
- Kant, Immanuel ([1781/1787] 2000): *Critique of Pure Reason*. Tr. and ed. W. S. Pluhar. Indianapolis: Hackett.
- Kant, Immanuel (1996): *Practical Philosophy*. Tr. and ed. M. J. Gregor. Cambridge: Cambridge University Press.
- Rohlf, Michael (2018): "Immanuel Kant." In: E. N. Zalta (Ed.): *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition). URL = <https://plato.stanford.edu/archives/sum2018/entries/kant/> (11–20–2019).
- Wood, Allen (1984): "Kant's Compatibilism." In: A. Wood (Ed.): *Self and Nature in Kant's Philosophy*. Ithaca, NY: Cornell University Press.
- Xie, Simon Shengjian (2009): "What Is Kant: A Compatibilist Or An Incompatibilist? A new interpretation of Kant's solution to the free will problem." In: *Kant-Studies*. 100(1), pp. 53–76.

Markus Kohl

Spontaneity and Contingency: Kant's Two Models of Rational Self-Determination

Abstract: I argue that Kant acknowledges two models of spontaneous self-determination that rational beings like us are capable of. The first model involves absolute unconditional necessity and excludes any form of contingency. The second model involves a form of contingency which entails alternative possibilities for determining oneself. Given these two models, the principle that we are exclusively determined by natural causes poses a twofold threat for human agency. In one respect (in relation to the second model), it threatens us with the obliteration of contingency, or with the universality of hypothetical necessity. But in another respect (in relation to the first model), it threatens us (and our putative “laws”) with the obliteration of absolute necessity, or with the universality of contingency.

1 Introduction

Leibniz argued that three conditions are individually necessary and jointly sufficient for freedom: intelligence, spontaneity, and contingency. Leibniz also claimed that since human minds are mirrors of God, human freedom satisfies these three conditions just like divine freedom (although to a different degree).¹ In this essay, I argue that for Kant intelligence and spontaneity are also necessary conditions of free agency. But in Kant, the relation between freedom and contingency differs for divine and human agency. Kant thus countenances two essentially different models of rational self-determination. The situation is further complicated by the fact that for Kant human agency does partake in the divine model, but only up to a very limited extent. If we appreciate that Kant has a twofold model of rational self-determination, we can differentiate between two very different senses in which Kant sees free agency threatened by the assumption that “all causality...[is] mere nature” (A534/B562).

My focus in this essay is on the rational self-determination that we achieve when we exercise our freedom of *will*.² Thus, my argument is that in Kant's view

¹ See Adams 1994, 11 and Jolley 2005, 5.

² Rational self-determination (or free, spontaneous agency) is not simply the same as the exercise of free will, because in Kant's view we also rationally determine ourselves when we exercise

our freedom of will has a twofold significance (which corresponds to the two sub-capacities that Kant distinguishes as conceptually separable components of our will, namely pure practical reason and Willkür) and that such freedom is threatened by the causality of nature in a twofold manner.

2 Intelligence and Spontaneity

At least in his ‘critical’ (1780s and 1790s) doctrine, Kant deems intelligence and spontaneity essential to freedom of will. In this section, I explain what these notions mean and why Kant requires that genuinely free agency must exemplify both these notions.

Consider first the notion of ‘intelligence,’ which is integral to the “standpoint” from which “a rational being [considers itself] as belonging to the intelligible world.”³ An intelligence “cannot think of the causality of its own will except under the idea of freedom.” When a rational being regards itself “as an intelligence,” it must regard itself “not from the side of its lower faculties” but from the side of its higher, empirically unconditioned faculties (4:452). Our *lower* faculties characterize us as merely passive beings; they include our capacity for receiving sense-impressions (including desires) and our empirical (reproductive) imagination. We have no active rational control over the exercise of those faculties: “The lower faculties cannot be instructed, because they are blind” (27:244). By contrast, our higher faculties *are* capable of being instructed, because the exercise of these faculties is a free activity governed by laws that contain normative standards of correctness: an intelligence “cognize[s] laws for the employment of its powers and all of its actions” (4:452). The normative laws governing the exercise of the higher faculties are “independent of nature, are not empirical, but are founded in reason alone” (4:452; cf. A546–547/B574–575).⁴

Consider next the notion of spontaneity. Kant distinguishes between a merely relative spontaneity (*spontaneitas secundum quid*) and an absolute spontaneity (*spontaneitas simpliciter talis*) (see, e.g., 5:96–101; 27:505; 28:268). The merely

our freedom of (theoretical) thought. See Kohl 2015a for a detailed analysis of Kant’s notion of doxastic freedom of thought.

³ For the analytic link between ‘intelligence’ and ‘intelligible’ or ‘noumenal being’, see e.g. 6:226; 28:583. For a helpful survey of how Kant uses these cognate terms, see Puls 2015, 184–185.

⁴ For a detailed argument that the non-empirical laws of reason are normative laws, and that Kant’s standpoint distinction is a contrast between the empirical and the normative point of view, see Kohl 2018.

relative notion – which Kant associates closely with the Leibnizian position (see, e.g., 27:505) – applies when actions are based on some inner principle that is ultimately not within the agent's control because it is implanted in (10:131; B167–168) the agent by an alien cause. Kant mockingly illustrates such relative spontaneity with artifacts such as turnspits or clocks, which, once designed and set up, operate according to inner mechanical principles. But it does not matter whether these inner principles are mechanical or psychological. We would have the mere “freedom of a turnspit” if our spontaneity was exhausted by our capacity to act on the basis of inner representational states that ultimately depend entirely on causes beyond our control (5:95–99; 27:503–504). Genuine freedom for Kant thus requires the spontaneity to act independently of the (however remote) determination by an alien cause. It requires the capacity for absolute *self*-determination: we act freely when “we are...determined by nobody, but determine ourselves and hence we possess spontaneity” (29:902–903).

Now, for Kant any form of agency is governed by a universal rule (9:11). Likewise, the concept of a cause and the concept of a universal law are analytically connected (B5; A539/B567; 4:446). Thus, any type of causally efficient activity that is *free* cannot be lawless either: lawless freedom is an absurdity. Since a genuinely free agent must be able to exercise her capacities without being determined by some alien cause, the law that governs the exercise of her capacities cannot be a law of nature: causal powers which are governed by laws of nature are always such that “something else determines the efficient cause to [its] causality” (4:446). The law that governs the causality of a genuinely spontaneous agent cannot be imposed on that agent by a *super-natural* designer or ‘puppeteer’ (5:101) either. Since a genuinely free, absolutely spontaneous activity must be independent of determination by alien causes, the law that governs such activity must spring from the agent's own rational nature: it must be a law of autonomy that the agent gives (legislates) to herself.

When Kant refers (at 4:452) to laws that an intelligence cognizes for the exercise of its higher faculties, he means precisely such laws of autonomy: self-legislated normative standards of good and bad agency. Every absolutely spontaneous agent is governed by autonomous laws that govern the proper exercise of the higher faculties which the agent possesses qua intelligence. Understood in a positive sense that designates not only the absence of determination by alien causes (this is the “negative” concept of freedom), spontaneity is the capacity to determine oneself through (laws of) reason (27:494). Having this capacity is necessary and sufficient for existing as an intelligence. Thus, for Kant the notions of an absolutely spontaneous agent and of an intelligence are co-extensive. An absolutely spontaneous agent that is a *practical* intelligence possesses the higher faculty

of a free *will* whose exercise is governed by an autonomous *moral* law that determines normative standards of morally good and bad agency.

3 Freedom and Contingency

Thus far, I have expounded Kant's view that intelligence and spontaneity are interconnected and necessary for free agency. But what about contingency, the third item in the Leibnizian trifecta? Here Kant's view raises important complications. To see why, consider the following two passages. The first is from the 1783 *Prolegomena to Any Future Metaphysics*:

The idea of freedom has its place solely in the relation of the intellectual, as cause, to the appearance, as effect. We [cannot] find any concept of freedom to fit a purely intelligible being, e.g., God, insofar as his action is immanent. For his action, although independent of causes determining it from outside, nevertheless is determined in his eternal reason, hence in the divine nature. Only if something should begin through an action, hence the effect be found in the time series, and so in the sensible world...does the question arise of whether...the concept of this causality is a concept of natural necessity... [or] of freedom. (4:344)

The second passage is from the 1793 *Religion Within the Bounds of Reason Alone*:

There is no difficulty at all in uniting the concept of freedom with the idea of God, as a necessary being: because freedom does not consist in the contingency of the action (that it is not determined through reasons), i.e., not in the indeterminism (that it be equally possible for God to do good or evil, if his action is to be named free), but in the absolute spontaneity, which faces a risk only with regard to predeterminism, where the determining ground of the action is in the preceding time, indeed in such a way that now the action is no longer within my power but in the hand of nature [which] irresistibly determines me; this difficulty now vanishes since in God no time series is to be thought. (6:50)

The first passage claims that we have no concept of freedom that fits a divine being because the actions of such a being would be entirely determined from within, "in his eternal reason." The second passage claims that it is entirely unproblematic to think of God as free because we can (indeed, must) think of divine agency as spontaneous rational self-determination. Thus, the first passage suggests while the second passage denies that contingency is essential to freedom.

It is not plausible to suggest that Kant changed his mind between 1783 and 1793. In Collins' notes on Kant's lectures on moral philosophy, which are usually dated to 1784–1785, Kant already affirms that divine actions can be both abso-

lutely necessary and free (27:267). In a series of handwritten notes that belong to roughly the same period, Kant first defines freedom (in the vein of the *Religion* passage) as the capacity to be only determined by reason (18:181) and then explains (more in the spirit of the *Prolegomena* passage) transcendental freedom as the “complete contingency of the actions” (18:183). I want to suggest that Kant is ambivalent here because he is trying to do justice to two different models of rational self-determination.

In section I, we saw that for Kant free agency essentially involves the capacity to act in accordance with autonomous laws of rational self-governance. Given this presupposition, the absence of contingency cannot count against free or spontaneous agency *per se*. A divine being would be inevitably determined by its eternal reason to do only what is right and good, and thus it would satisfy the conditions of spontaneous, intelligent self-determination *par excellence*. Divine freedom is the highest kind of freedom because it involves a ‘moral necessity’ which arises from absolute spontaneity (28:806; 28:1068; 28:1280–1281). This is the point of the *Religion* passage.⁵

But now consider a type of being that, like a divine being, has the capacity to act in accordance with self-legislated objective standards of right reason but that, unlike a divine being, is also (and necessarily so) under the influence of subjective, non-rational sensible conditions. Here radical contingency is a constitutive feature of free agency:

But if reason of itself does not sufficiently determine the will, and if the will is subjugated to subjective conditions (certain incentives) which do not always agree with the objective conditions – in a word, if the will is not of itself in accord with reason (which is really the case with human beings), then the actions which are recognized as objectively necessary are

5 There is a complication. In the *Prolegomena* Kant says that we cannot find a concept of freedom for divine action “insofar as that action is immanent.” One might suggest that the *Prolegomena* thus allows us to regard as free the non-immanent divine act of creating the world. Indeed, in student notes on Kant’s lectures one can find the idea that God creates the world according to His freedom, in contrast to His inner actions that follow from the necessity of His inner nature (28:1298–1299; 1092–1097). However, God’s creation of the world cannot satisfy the *Prolegomena* definition according to which a free action must have a sensible effect that begins in the time series: for we cannot regard the sensible world as coming into existence at a certain temporal moment, a first beginning of the time series (cf. A427/B455 ff.). Moreover, we are still left with the conflict between the claim (1) that divine freedom requires the absence of rational determination from within, and the claim (2) that divine freedom allows for such determination. The *Religion* passage provides a (to my mind) strong philosophical reason in favor of (2). If *any* divine action, including His creation of the world, were contingent, this would entail that in this action God is *not* absolutely determined by standards of right reason: in that case, the action (and its result) seems rationally arbitrary, which is incompatible with the idea of divine perfection.

subjectively contingent [...]. That is, the relation of objective laws to a will which is not completely good is conceived as the determination of a rational being by principles of reason to which this will is not by its nature necessarily obedient. (4:412–413)

There is a complication here: Kant denies that contingency can enter into the definition of free agency. For Kant, a real definition must demonstrate the “objective reality” of a thing (A241–242). This requires that only positive determinations rather than mere negations be included in the definition, since “negation signifies a mere want, and, so far as it alone is thought, represents the abrogation of all thinghood” (A575/B603). Freedom positively consists in the power to act in accordance with self-legislated rational laws. Hence, the possibility to deviate from rational laws is not a positive capacity of its own but a mere want of a capacity, a negation or “privation” that as such cannot figure in a definition of free agency (6:226–227). But nonetheless, the possibility of deviation can be a necessary component of free agency, if the agent in question is *necessarily* under the influence of conditions that entail this possibility. This is true for human agents. As rational creatures, we always have the power to determine ourselves in accordance with the normative standards of our own reason, but we are also always subject to non-rational influences that account for a pervasive tendency to violate these standards. While the possibility to deviate from standards of the (morally) right and good qua privation cannot figure in a definition of human freedom, it nevertheless attaches necessarily to our free agency as the perennial ontological ‘shadow’ of our rational power and thereby afflicts our free agency with a radical form of rational contingency.⁶

Now the *Prolegomena* passage does not explicitly invoke our rational imperfection as the reason why the concept of freedom that we deem suitable for ourselves does not apply to divine agency. But it does mention a closely related feature: it cites the “immanence” of God’s actions as the reason why we cannot apply our concept of freedom to such actions, and it contrasts the immanence of divine actions qua “determined by His eternal reason” with the fact that our free agency is supposed to have effects in the sensible world. That is, our free agency (insofar as it involves the will) is always directed towards some ex-

⁶ I thus cannot agree with those (like Pereboom 2005) who deny that for Kant free agency requires alternative possibilities. While this is true for divine freedom, the freedom of imperfectly rational agents is essentially free agency under categorical *oughts* that represent the rational necessity (and therefore the possibility) of acting according to a rule, whose transgression always remains a possibility as well (29:1017). The very concept of an objective imperative of reason presupposes the possibility of satisfying the command as well as the opposite possibility; this is incompatible with predeterminism (29:1019). (See Kohl 2015b for a more detailed discussion of this issue.)

ternal *telos* that lies outside the free exercise of our rational capacities properly speaking: even when we have done everything that is fully within our rational control, i.e., when we have chosen the right course of actions for the right reasons, there is still a question as to whether we will succeed in performing the chosen action or in producing the intended effects (cf. 4:394). Whether we act in the way that we intend and whether we effect the changes that we intend depends upon vagaries of the empirical world such as our physical state. These vagaries are always to some extent unforeseeable and beyond our immediate rational control. (For example: I might be firmly committed, for the right reasons, to execute my promise to meet a student on campus, but when I get going, I faint and collapse.) This circumstance affects our free agency with an additional layer of contingency: beyond the fact that it is contingent (because it is up to us) whether or not we choose for the right reasons, it is also contingent (because it is not completely up to us) whether or not we succeed in realizing our chosen ends.⁷

The fact that our rational (volitional) self-determination always refers to a *telos* outside itself is yet another consequence of the fact that we are essentially rational *creatures*: as such we lack the absolute self-sufficiency that is the privilege of a divine being, and our agency must always be directed towards some concrete end which we strive to realize in the sensible world. For Kant, the lack of immanence in our moral agency is systematically connected to our rational imperfection: both features derive from our sensible nature. The fundamental moral prescriptions that Kant classifies as narrow duties regulate our pursuit of non-moral ends, i.e., they prescribe the conditions under which we can legitimately attempt to realize certain empirical states of affairs that appeal to us given our non-moral, sensible needs and desires. Narrow duties set limits to our tendency to prioritize our non-moral ends (chiefly, our wish to be happy) over the moral value of dignity or personality (including our own personality). But even when morality positively commands us to pursue certain ends, these ends depend on our constitution as imperfectly rational creatures. The moral directive to promote the highest good takes into account our need to combine our moral vocation with our sensible desire for happiness. Regarding the more mundane external moral ends that we pursue under the heading of our wide or imperfect duties (e.g. the end to promote the happiness of others or to cultivate our talents), Kant holds that if we were not subject to sensible incentives that attract us to the pursuit of immoral ends, we would not need the countervailing influ-

7 I discuss the significance of this point for Kant's verdict that *ought implies can* in Kohl 2015c.

ence of the external (“material”) moral ends that are prescribed by our imperfect moral duties:

For, since the sensuous incentives incite towards ends (as the matter of the faculty of choice) which may be contrary to duty, the law-giving reason can fend off their influence in no other way than, again, through an opposing moral end [...]. (6:380–381)

Conversely, the immanent actions of a self-sufficient being need not be directed towards ends or effects that lie outside the internal sphere of its rational self-constitution.⁸

For Kant, we only have a highly abstract, indeterminate notion of absolutely self-sufficient, perfectly rational agency. Concepts that are essential to the structure of human (or more generally finitely rational) agency, such as “imperative,” “incentive,” “interest,” or “maxim,” are all inapplicable to a divine being (5:79). We can ascribe a will to God only symbolically, by means of analogy, because our literal conception of a will is tied to our awareness of limitations, such as the dependency of our contentment on external objects that (partly) explains why we strive to realize these objects; if we abstract away from those limitations, the very concept of will is also eliminated (28:780–782; 798).

Kant also faces difficult questions about whether our concepts of causality or action are really suitable for the representation of a divine being – questions that may ultimately push Kant to the admission that we have only a non-literal, analogical or symbolic way of representing a divine being as an agent or as a cause (see 5:464–465; 482–485).⁹ Hence, we might take the *Prolegomena* passage to

⁸ Here we must again (see footnote 5) note the complications arising from the non-immanent divine act of creation, which does seem to be directed towards an external end. Student lecture notes show Kant wrestling with this problem. The main question here is: why does God, as an absolutely self-sufficient being, create something outside of Himself (28:1060)? It seems that in order to comprehend this, we must attribute to God some kind of interest or incentive, but Kant rejects this implication as inconsistent with God’s absolute perfection and infinitude (28:780–782; 1060; 1065; 1201–1204; 1279). Kant sometimes suggests that we must attribute to God something analogous to an interest (28:1279), but it is unclear what this amounts to. Another suggestion he considers is that God’s motive for creating things outside of himself is his self-sufficiency, namely, the non-dependent pleasure deriving from his self-consciousness (1) as a self-sufficient ground of all possibility and (2) as possessing the power to produce the best of all possibility (28:1061–1062; 1100–1102; 1204; 1257–1277). I confess that I do not fully understand this proposal, but at least it has the virtue of nominally tying God’s external causality to His immanent, self-sufficient rational nature.

⁹ Part of the problem here is that our conceptual resources seem generally unsuited for thinking any intelligible being in more than an analogical fashion (cf. A566/B594). I discuss this issue in

mean that we cannot find a *determinate* concept of freedom that is suitable for a divine being. We do have a very abstract, even minimally positive way of representing a perfectly rational being as free, namely, through a notion that designates only God's absolutely necessary conformity with self-legislated standards of right reason. But the kind of free causality that we recognize in ourselves, from our lived deliberative perspective, is of a vastly different kind. Our free agency is governed by autonomous norms of right reason, and it is always a matter of absolutely spontaneous self-determination, which accounts for some continuity with divine freedom. But as creatures who are affected with sensible needs and desires, we are constantly attracted to choosing on behalf of our sensible as opposed to our rational nature (on behalf of happiness as opposed to morality); moreover, our free agency always involves a relation to external ends that we can realize only if empirical conditions beyond our immediate control happen to cooperate. Thus, our free agency unlike divine freedom involves two (connected) layers of contingency.

4 Two Types of Human Freedom

However, there is a further complication, which again can be uncovered by considering two conflicting sets of passages. In the *Critique of Pure Reason*, Kant refers to "the actions through which [reason] prescribes laws," which must be considered free in a practical sense (A803/B831). Through these actions, reason "yield[s] us laws that are pure and determined completely a priori [...] which are prescribed to us not in an empirically conditioned but in an absolute manner" – moral laws as "products of pure reason." (A800/B828) In its prescriptive activity, reason "frames for itself with absolute spontaneity an order of its own according to ideas [...] according to which it declares actions to be necessary [...]" (A548/B576).

By contrast, consider this famous passage from the 1797 *Metaphysics of Morals*:

From the will spring laws, from the Willkür spring maxims. The latter is in human beings a free Willkür; the will, which concerns nothing other than merely the law cannot be called either free or unfree, because it is concerned not with actions but immediately with the legislation for the maxims of actions [...] and which is therefore absolutely necessary and capable of no necessitation. Hence only the Willkür can be called free. (6:226)

Kohl 2015d. But the problem is, I believe, especially acute with regard to God as an intelligible entity, as I argue in Kohl (forthcoming).

Thus, in the first *Critique* Kant says that moral laws are products of the free, absolutely spontaneous prescriptive activity of reason, whereas in the *Metaphysics of Morals* he denies that the concept of free action can be applied to the will or practical reason as a law-giving faculty.

There is room to doubt that the second passage expresses Kant's considered view. In his notes leading towards the *Metaphysics of Morals* (*Vorarbeiten*), we find Kant wavering, at one point (23:248) denying and at another point (23:383) affirming that the legislative will is free in the negative sense that denotes the absence of determination by foreign causes. Moreover, if he did unequivocally deny the possibility that we can speak of a free act of practical legislation, this might have dire consequences for his overall account.¹⁰ One casualty would (arguably) be the 'fact of reason' doctrine of the second *Critique*. For in this doctrine, Kant argues that the moral law is given by a legislative deed of reason, a *Tathandlung* that is independent of determination by foreign causes such as empirical conditions or divine imposition; this seems to imply that the act of prescriptive judgment that provides the relevant law must be a free action. Accordingly, in the *Groundwork* Kant says that the freedom of reason consists in not "receiving a bias from any other quarter with respect to its judgments," and consequently in the fact that reason can "regard itself as the author of its principles independent of foreign influences" (4:448).

Therefore, I hesitate to accept the *Metaphysics of Morals* passage as expressing Kant's considered view. But I think there is a way of making good sense of the spirit, if not the letter, of what Kant says in this passage. I believe that Kant's ambivalence concerning the idea that the legislative act of practical reason may be called a 'free action' is quite closely related to his ambivalence concerning the idea that a divine being may be called a 'free cause'. In both cases, his denial is driven by the recognition that our ordinary notion of free agency indeed does not fully *suit* these cases: neither divine agency nor practical legislation instantiate quite the same type of freedom that we attribute to ourselves in the deliberative acts of choice that serve as the paradigm for our 'experience' of freedom. Practical legislation, like divine agency, is an entirely "immanent" action: it is not directed towards some further course of action or the pursuit of some external end. Rather, practical legislation is complete in itself, driven solely by its own inner logic, to the representation of what is the practically necessary or objectively good course of action: it "concerns nothing other than merely

¹⁰ These consequences are overlooked by those (like Beck 1960, 172, 185 and Bojanowski 2006, 242–244, 261) who insist that practical reason has no effects apart from its power to determine Willkür (in morally good choices).

the law" (6:226) itself. Moreover, this inner logic is as irresistible and inevitably determining for our practical thought as the force of rationality is without exception for *all* divine agency. We cannot struggle against the rational activity that imposes practical laws upon our will, i.e. we cannot choose not to be governed by such laws: even immoral choices are informed by a however dim or twisted awareness of the supreme rational authority that the moral law has over our will. The moral law is recognized as an apodictically certain, indubitable fact of reason (5:29–31, 47; 6:225), and hence we cannot refuse to recognize that morality has a rational claim on us even when we are inclined to transgress its commands: the moral law "forces itself upon us" (5:31) "whatever inclination may say to the contrary" (5:32); the voice of reason here is "irrepressible" (5:35). In acts of practical self-legislation, we are inevitably determined by reason.

Of course, the inevitable determination by reason ends for finite agents like us with the mere consciousness of what would be the (morally) right and good thing to do; for us the determination by reason does not automatically lead towards acting in accordance with this consciousness. We need to be rationally *necessitated* to actually choose what is right and good. The concept of rational 'necessitation' tracks the determining force of reason in relation to internal constraints or hindrances that inveigh against reason. The presence of such hindrances crucially affects our acts of free choice, but it does *not* affect the act of legislation itself. Hence the notion of necessitation is inapplicable here: "the legislation for the maxims of actions [...] is [...] absolutely necessary and capable of no necessitation" (6:226). And this is precisely why it may seem odd to speak of 'free action' with regard to the legislative deed of practical reason. This immanent, rationally inevitable deed does not fit the model of freedom that we are most intimately familiar with when we deliberate about which external ends we ought to pursue.

Thus, Kant's claim (at 6:226) that the concept of freedom applies only to our Willkür should be understood as pointing out that the legislative activity of our will or practical reason cannot be regarded as free in that precise sense which is most familiar to us from our ordinary deliberative standpoint, and which is also central to Kant when he defines "preliminary concepts for the Metaphysics of Morals." Here Kant seeks to delineate a concept of free action or "deed" that suits our notion of moral and legal accountability. The acts or deeds for which we consider ourselves (and others) responsible are those that are subject to our Willkür (6:223). The act of legislative reason that provides us with a consciousness of the moral law is a necessary *precondition* for our moral and legal accountability, but it is our choices and our intentional observable (physical) actions, rather than the act of practical legislation itself, that call for praise, blame or punishment.

Kant's conception of practical legislation entails that we partake in the divine model of free agency, albeit only to a strictly limited extent. A divine being has no separate faculty of Willkür for the choice of maxims. It is determined by reason with absolutely necessity all the way through, and capable of no necessitation whatsoever. In our case, the legislative will as practical reason does determine us, or better: we rationally *determine ourselves* to the absolutely necessary legislative deed that yields the moral law, but this is where the inevitably determining influence of reason ends and the need for a necessitating influence by reason begins. Our execution of the normative standards which we represent in the idea of the moral law hinges on our exercise of a separate volitional capacity, the capacity for choice or Willkür, which is subject to conditions that inveigh against the moral law.¹¹

Kant has another piece of terminology for capturing this difference. In our acts of practical legislation, our will or practical reason functions as “pure reason,” as a “capacity of principles (and, here, of practical principles, indeed as a law-giving capacity)” (6:214); by contrast, in our acts of practical choice our Willkür is “affected, but not determined through [sensuous] incentives, and is thus in itself [...] not pure but can be determined through actions from a pure will” (6:213) – namely, when we choose to determine ourselves in accordance with the pure self-legislated practical (moral) law. The legislation of practical reason is “pure” because it is *not even affected* by a countervailing influence of sensibility, whereas the choices of our Willkür are free from determination but not from affection by sensuous incentives (cf. A534/B562). This affection makes us apprehend the pure practical laws as a command, imperative or ought.¹² Kant characterizes a divine will as entirely “pure” because a divine will is never affected by any other things. Accordingly, a divine being does not apprehend the pure rational law as an ought, it does not have a faculty of Willkür, and it does not have the kind of freedom that is characteristic of this faculty: a freedom that, while not *definable* as a liberty of indifference, nevertheless es-

¹¹ I do not wish to posit will and Willkür as *ontologically* separate faculties. Kant indicates that there is a *conceptual and a functional* difference between the two when he says: “Laws arise from the will, viewed generally as practical reason; maxims spring from Willkür.” (6:226). But this does not imply that we could have a will in isolation from a Willkür or vice versa. Here I agree with Allison 1990, 130–131 and Beck 1960, 190. They both emphasize that will (in the narrow, legislative sense) and Willkür are two interconnected aspects of one faculty (‘will’ in a wide sense). I briefly return to this point in the conclusion.

¹² Compare 4:457–458, where Kant also distinguishes pure (sensibly unaffected) reason that “gives” (i. e. is the source of the representation of) the moral law from the faculty that is responsible for choice, whose spontaneity is affected (but not determined or necessitated) by sensible inclination.

entially involves an element of arbitrariness ('willkürlich' just means 'arbitrary') that it inherits from the imperfectly rational character of the sensibly affected faculty of choice to which it attaches.

5 Two threats that nature poses to human freedom

I have argued that Kant canvasses two distinct models of free, rational self-determination. In each model, absolute spontaneity and intelligence are necessary conditions of free agency. In the first model, the absolutely spontaneous intelligence is inevitably determined by reason to a certain act. In the second model, the absolutely spontaneous intelligence is affected by sensible conditions whose influence inveighs against reason, which makes it contingent whether or not the agent acts in accordance with right reason. A divine being would instantiate the first model in all of its actions. We partake in this divine model in our pure, sensibly unaffected act of rational legislation. We instantiate the second model in our sensibly affected free acts of choice.

I suggest that we use the terms 'legislative freedom' and 'executive freedom' for the two kinds of freedom that we exhibit.¹³ In the case of a divine being, this distinction is superfluous as there is no conceptual space for a gap between absolutely necessary legislation of practical laws and absolutely necessary execution of the rational standards that are cognized through these laws. Since both terms designate absolutely spontaneous acts that are not determined by alien causes, both executive and legislative freedom can be regarded as species of *transcendental freedom* in the *negative sense* of that term. In my view, we can also regard both executive and legislative freedom as falling under Kant's *positive* concept of transcendental freedom, though there is some debate about the proper extension of that concept.¹⁴

¹³ This distinction is not, to my knowledge, already drawn explicitly in the literature on Kant's doctrine of freedom, though I believe that it is compatible with seminal interpretations such as Allison 1990 or Beck 1960.

¹⁴ It is beyond dispute (accepted by e.g. Allison 1990; Beck 1960; Bojanowski 2006; Willaschek 1992) that *Willkür's* choice of a morally *bad* maxim or action is not an act of autonomy or positive freedom. Such an act is free in the negative sense (not determined by alien causes), and it is still *governed by* ("stands under") the positive law of autonomy; these two points secure that the agent is accountable for the immoral choice (see Bojanowski 2006, p. 257). The dispute arises over the following issue: does the positive concept of freedom as autonomy apply (1) to the will qua legislator of laws (Beck 1960, 199–200; cf. 5:33); or (2) to the will in the wide sense

The main reason why the distinction between executive and legislative freedom strikes me as important is that it allows us to distinguish between two different threats that arise, in Kant's view, from the supposition that "all causality... [is] mere nature" (A534/B562). To adopt this supposition is "to make principles of possible experience conditions of the possibility of things in general" (A781/B809). One such condition is the principle that everything is *hypothetically necessary*, i.e., that everything is the necessary result of a preceding causally efficient state which in turn results from some further preceding cause and so on and so forth, without a first or final, causally unconditioned member of the chain that would be *absolutely* rather than hypothetically necessary. Hence, the upshot of turning the principles of possible experience into absolute metaphysical conditions is this: hypothetical necessity is the only non-trivial (non-logical) form of necessity that exists; *everything* (every state and event) is the result of hypothetical necessity; *all* efficient causes operate according to the principle of hypothetical necessity (natural causality), which means that their causality is dependent upon some further cause (cf. A542/B570). Call these *naturalistic* suppositions.¹⁵ Under these naturalistic suppositions, both our legislative freedom and our executive freedom would be impossible, but for very different reasons.

Naturalistic suppositions (if they were true) would rule out the possibility of our executive freedom because they would remove the absolute contingency that is characteristic of imperfectly rational agency under laws of reason. If all our causal capacities operate in accordance with the principle of natural, material or hypothetical necessity (for this terminology, see A226/B279-A230/B283), then we are always inevitably determined by some temporally preceding natural condition (or set of conditions) to exercise these capacities in one particular way, namely to perform a particular act of choice at a given temporal moment (cf. A534/B562). In this scenario, we could not be governed by laws of practical reason in the only way that is possible for imperfectly rational agents like us, namely, via categorical imperatives (*oughts*). In order to be governed by such imperatives, agents must have the option to go either way with regard to laws of reason, because they must have the power to act in accordance with right reason as well

(as including both legislative will and Willkür; Allison 1990, 132); or (3) to the actual determination of Willkür by the legislative will (Bojanowski 2006, 256–257)? Since (1)–(3) are so intimately connected, it seems to me reasonable to conclude that (1)–(3) can *all* be regarded as positively free. I do not see that any substantive philosophical or interpretive issue hinges on picking one of these options at the exclusion of the other two.

¹⁵ A label that better accords with Kant's own usage might be 'suppositions of dogmatic empiricism' (cf. A466/B494; A471/B499).

as the option to deviate from right reasons (an option that derives from our perennial shadow of rational imperfection).¹⁶

To be sure, if all our actions were exclusively governed by causes that operate according to the principle of natural causality, then our actions (as natural changes) would be contingent in a relative or empirical sense that denotes their dependency on some temporally prior causal condition (cf. A460/B488; B290 – 291). But this sort of contingency cannot save our self-image as free, accountable agents, for it means that the causal condition (or the set of such conditions) that renders a particular action hypothetically necessary is itself hypothetically necessary under some further causal condition, and so on and so forth until we reach a point in time where the relevant condition is no longer under our rational control. This situation is what Kant refers to as “predeterminism” in the *Religion* passage cited above (6:50).

The above suppositions would also rule out our *legislative* freedom and thereby the existence of laws of autonomy that derive from pure acts of legislation. This is because for Kant such laws purport to prescribe with unconditional rational necessity how we must act; so the legislative acts that yield such laws, i.e. the “actions through which [reason] prescribes laws” (A803/B831), must also be considered “absolutely necessary” (6:226). Morals laws “as products of pure reason” “are prescribed to us not in an empirically conditioned but in an absolute manner” (A800/B828). This would be impossible if all our actions were exclusively subject to hypothetical necessity. In the natural order of things, there is only the empirically conditioned (hypothetical, material) necessity of actions or states relative to some temporally preceding condition (cf. A228/B280); everything in nature is empirically conditioned and therefore contingent (cf. A560/B588). Hence, if there were only the natural order of things, or (alternatively) if everything were subject to the conditions that govern the natural order of things, then *no action and no product of any action* would be absolutely or unconditionally necessary: our (purported) acts of legislation and their products, i.e. the laws that spring from these acts, would always be contingent upon empirical conditions outside of our intellectual self-control, such as our social conditioning, our gene pool, our empirical psychology, etc. Since judgments that purport to represent rational laws lay claim to a strict a priori necessity, such judgments would be invalid, mere “figments of the brain,” if they were the upshot of contingent empirical conditions. This line of reasoning is absolutely central to Kant's thinking:

16 In Kohl 2015b I provide for a systematic argument for this claim.

If the determining grounds were empirical and were given in an a posteriori subjective fashion, then the judgment of reason could not be regarded as a priori and thus could not be regarded as absolutely necessary. In order to judge in objectively universal, and indeed in an apodictic fashion, reason must be free from subjectively determining grounds; for if those grounds did determine [reason], then the judgment would be merely as it is contingent, i.e., according to its subjective causes. Hence, reason is conscious of its freedom in objectively necessary judgments a priori [...]. (18:176)

...the understanding alone (and the will, insofar as it can be determined through the understanding) is free and pure self-activity which is determined through nothing other than itself. Without this original and immutable spontaneity, we would cognize nothing a priori; because we would be determined to everything, and our thoughts themselves would stand under empirical laws [...]. (18:182–183)

The title to freedom of the will claimed by ordinary reason is based on the consciousness and the conceded presupposition of the independence of reason from merely subjectively determining causes which together constitute what belongs only to sensation and is included under the general name of sensibility. (4:457)

[Even the fatalist always presupposes that] the understanding has the capacity to determine its judgment according to objective reasons that are valid at any time, and does not stand under the mechanism of merely subjectively determining causes that may change in the meantime (8:14)

If [...] a judgment [...] asserts a claim to necessity, then [...] it would be absurd to justify it by explaining the origin of the judgment psychologically. For [...] if the attempted explanation were completely successful it would prove that the judgment could make absolutely no claim to necessity, precisely because its empirical origin can be demonstrated. (20:238)

Kant's point here can be summarized as follows. Suppose we had to concede that our representation and acceptance of what we regard as our fundamental practical principles is the upshot of contingent, non-rational empirical conditions, such as our empirically given desires or our gene pool in conjunction with our natural environment. For Kant, the concession that we are determined in our practical thinking by contingent, non-rational, and mutable causal circumstances is incompatible with what he regards as an essential presupposition of our stance as moral agents: namely, the presupposition that our fundamental practical principles have a timeless validity which allows them to make an unconditional rational claim on all rational agents regardless of what their empirical conditions happen to be and however these conditions might change in the future. Since these principles can only exist as laws of autonomy that derive from our own rationality, the timeless validity of these principles presupposes that these principles derive from the timeless, "original and immutable spontaneity" of our practical reason, which is free from "the mechanism of subjectively determining causes": causes that, qua empirical, are as contingent and mutable as their products or effects. In short, if all our acts of judgment were determined

by contingent empirical conditions, then our rational faculties would lack the authority to prescribe with strict necessity how we must act regardless of what our empirical conditions happen to be.

If my interpretation here is correct, then for Kant the assumption that all our actions are exclusively subject to natural causality raises two conceptually different worries for human freedom. Only the first of these worries hinges on the fact that our executive freedom involves metaphysical contingency or alternative possibilities. Our actions can be governed by categorical imperatives, and we can be morally praise- and blameworthy for our actions, only if we can go either way with regard to such imperatives; for Kant this would be impossible if all our actions were the inevitable result of (i.e. were exclusively determined by) an indefinitely extending chain of natural causes. But with respect to our legislative freedom, Kant insists that we are inevitably determined by our own rationality to perform the legislative act that provides our fundamental practical consciousness of the a priori moral law. Here the problem is not that the mechanism of nature does not allow for absolute contingency, but, rather, that this mechanism *qua* constituted by merely contingent empirical conditions and hypothetical necessities does not allow for the absolute necessity and for the timeless, immutable validity that we represent in our idea of a moral law of autonomy. Thus, “if all causality [were] mere nature,” then neither the absolute contingency nor the absolute necessity that we require, respectively, for our executive human freedom and for our legislative human freedom could exist. (If *everything* were subject to natural causality, then God as a being that acts freely outside of nature and time would of course also be impossible.)

I want to end with two clarifications. First, although I have presented these worries as pertaining to the supposition that ‘all causality is mere nature’, we have already seen (in section I) that nothing hinges on the supposition that all our actions are determined by an alien *natural* causality. Of course, if we assume that our actions are determined by the super-natural causality of a divine being, then there is no problem concerning the freedom of *that being*. The divine being in this scenario has the absolute freedom to act in accordance with the unconditionally necessary laws of its immutable, eternal rational essence. But since the laws that determine *our* actions in this scenario are imposed on us by an alien supernatural cause, *we* lack the legislative freedom to give rational laws of autonomy to ourselves. We also lack the executive freedom to act in accordance with or against these laws: our executive choices are hypothetically necessary upon the divinity’s creative imposition of the inner principle governing our causality, and hence our choices lack absolute contingency. In this scenario, the chain of hypothetically necessary causal conditions terminates in an absolutely

necessary first divine action that is not determined by some prior cause, so transcendental freedom does exist – but it does not exist in us.

Second, the two worries that I have considered are intimately connected: the worry concerning legislative freedom is logically prior to the worry concerning executive freedom. The latter worry can arise only under the supposition that there *are* rationally necessary laws that govern our choices, laws in relation to which we have both the power of compliance and the option of deviation. Hence, the second worry about whether we possess executive freedom presupposes that the first worry, about whether we possess the legislative freedom required for practical legislation, can be resolved. Now, if the first worry about legislative freedom can be resolved, this entails that the second worry about executive freedom can be resolved as well. If we really do possess (pace the first worry) the legislative freedom to prescribe *valid* laws of reason, this entails (pace the second worry) that we also possess the executive power to satisfy these laws: a will cannot (on pain of compromising its legislative authority) legislate normative laws to itself if it lacks the power to comply with these laws.¹⁷ If we do possess the executive freedom to comply with laws of reason, then our option to deviate from these laws simply follows along as an ontological limitation that is parasitic on our power of rational choice. Legislative freedom thus entails executive freedom.

But as long as we keep in mind that there are these deep interconnections between the two types of human freedom and (accordingly) between the naturalistic worries that one can raise for each type, it is profitable to discuss these worries separately so as to avoid confusion about what is at stake in each case. Kant shares the contemporary sense that the issue of (pre-)determinism versus alternative possibilities is central to human executive freedom, the kind of freedom that characterizes our faculty of Willkür and that we exercise in acts of choice. This is the first of his concerns about the implications that an exclusively naturalistic world has for human freedom of will, a concern which depends on the assumption that the natural order of things is a deterministic order. But Kant also has a conceptually separate second concern about the implications of an exclusively naturalistic worldview, which pertains to the contingency of all human thought and judgment: a concern which does not hinge on the (from a modern standpoint controversial) issue of whether the contingent natural order of things is conceived as deterministic or indeterministic. This second concern arises for the kind of freedom that characterizes our faculty of will (in the narrow sense, qua pure practical reason) which we exercise in acts of legislation,

17 I expound this point at greater length in Kohl 2015c.

when we arrive at (purportedly) a priori necessary practical judgments about what our moral duties or obligations are.

My suggestion that Kant raises two interconnected yet conceptually separate worries about the implications that a naturalistic worldview has for our freedom of will neatly corresponds to Kant's view that the human will (in the broad sense) is constituted by two sub-capacities that are ontologically intertwined but conceptually (or functionally) separable. The free will (in the broad, inclusive sense) of imperfectly rational agents comprises both a pure legislative faculty of practical reason (which Kant calls the will in a narrow sense) and a faculty of choice. For such a free will to exist, two metaphysical conditions must be satisfied. First, there must be absolute rational necessities, namely, prescriptive ideas of pure practical reason that derive from pure acts of moral legislation which are not contingent upon rationally arbitrary empirical conditions. Second, there must be absolute contingencies, namely, causally unconditioned acts of Willkür through which we can either choose morality over happiness or choose happiness over morality. Neither absolute rational necessities nor absolute contingencies can exist in the natural order of things, in the world as it is understood by contemporary naturalists. Hence, there can be no free will (or moral agency) on a naturalistic worldview.¹⁸

Bibliography

Quotations from Kant's works, apart from the *Critique of Pure Reason* which is cited according to the standard A/B pagination, cite the volume and page number of the Academy edition, *Kants Gesammelte Schriften*, ed. Königlich Preussische Akademie der Wissenschaften, vols. 1–29. Berlin: De Gruyter, 1900–. Translations are based on the Cambridge Edition of the works of Kant.

Adams, Robert Merrihew (1994): *Leibniz: Determinist, Theist, Idealist*. Oxford: Oxford University Press.

Allison, Henry (1990): *Kant's Theory of Freedom*. Cambridge: Cambridge University Press.

Beck, Lewis White (1960): *A Commentary on Kant's "Critique of Practical Reason"*. Chicago: University of Chicago Press.

Bojanowski, Jochen (2006): *Kants Theorie der Freiheit*. Berlin: De Gruyter.

Jolley, Nicholas (2005): *Leibniz*. London: Routledge.

Kohl, Markus (2015a): "Kant on Freedom of Empirical Thought". In: *Journal of the History of Philosophy* 53 (2), pp. 301–326.

¹⁸ For helpful comments, I am grateful to Stephanie Basakis, the audience at the 2017 APA Central Division Symposium on Spontaneity and Rationality, and the audience at the 2017 conference on the Concept of Will in Classical German Philosophy (LMU Munich).

- Kohl, Markus (2015b): "Kant on Determinism and the Categorical Imperative". In: *Ethics* 125 (2), pp. 331–356.
- Kohl, Markus (2015c): "Kant and 'Ought Implies Can'". In: *Philosophical Quarterly* 65 (261), pp. 690–710.
- Kohl, Markus (2015d): "Kant on the Inapplicability of the Categories to Things in Themselves." In: *British Journal for the History of Philosophy* 23 (1), pp. 90–114.
- Kohl, Markus (2018): "Kant's Standpoint Distinction". In: *Kantian Review* 23 (2), pp. 229–255.
- Kohl, Markus (forthcoming): "Die radikale Unbegreiflichkeit von Gott für den menschlichen Verstand." In: Bernd Dörflinger, Heiner Klemme (Eds.): *Die Gottesidee in Kants theoretischer und praktischer Philosophie, Studien und Materialien zur Geschichte der Philosophie*. Hildesheim: Georg Olms.
- Pereboom, Derk (2005): "Kant on Transcendental Freedom". In: *Philosophy and Phenomenological Research* 73 (3), pp. 537–67.
- Puls, Heiko (2015): "Kant über die äusserste Grenze aller praktischen Philosophie – Ein Kommentar zur Sektion 5 der Grundlegung". In: Dieter Schönecker (Ed.): *Kants Begründung von Freiheit und Moral in Grundlegung III*. Münster: Mentis, pp. 157–214.
- Willaschek, Marcus (1992): *Praktische Vernunft. Handlungstheorie und Moralbegründung*. Stuttgart: J.B. Metzler.

Halla Kim

How is the Corruption of the Will Possible? Kant on Natural Dialectic and Radical Evil¹

Abstract: In this chapter, I critically discuss Kant's account of the will in the context of our moral failure as he discusses its importance as part of his view of a 'natural dialectic' of the *Groundwork* and also as part of his well-known doctrine of radical evil in the *Religion*, respectively. In particular, I suggest that, even though Kant's understanding of the way our will fails to be moral in the natural dialectic may be a precursor to a more sophisticated account of will as the latter is developed in his doctrine of radical evil in the *Religion*, the natural dialectic is not quite "equivalent" to the radical evil. In Section 1, I will introduce the doctrine of natural dialectic in the *Groundwork* and show how the Kantian conception of will together with its maxim is deployed in the process of a natural dialectic. In Section 2, I will discuss in detail how Kant develops a sophisticated conception of will in his elaboration of radical evil in the *Religion*. Section 3 then suggests that, while the *Groundwork* account of the corruption of the will proceeds by means of the fallacy of subreption predominantly committed at the level of individuals, in the *Religion* Kant gives an account of the corruption of the will by means of the two-tier structure of will, that is, the distinction between the *Wille* and *Willkür* at the level of the human species as a whole. Section 4 attempts to shed light on the exact relation between the two accounts of the will. In particular, I will argue that, *pace* Henry Allison, they are not "virtually equivalent." Section 5 is a brief conclusion.

1 Introduction

It is well known that Schelling admired Kant for his insight to pinpoint moral evil in an act of freedom, while criticizing Fichte for falling prey once again to the philanthropism prevalent in his moral theory and for wanting to locate the origin of the evil that precedes all empirical action in the lethargy of human nature (cf.

¹ For abbreviations, please refer to the glossary provided. References to Kant's works shall also be given internally by the volume and page number of the Akademie edition (the standard edition of Kant's works in German) with the exception of references to the *Critique of Pure Reason*, which will be given, according to custom, by the pagination of the first ("A") and second ("B") editions.

Schelling 1887 [1807], 67). According to Schelling, Fichte lost his nerve in §16 of *System of Ethics*, ending up explaining moral evil in us by means of natural forces. Kant was right to attribute evil to an exercise of the will that is free.

Indeed, there is no question, the Kantian ethics of autonomy crucially depends on a conception of the will (*Wille*) together with freedom as its property—yet, it is not as well-known that the same will also plays a large role in our innate tendency to fail in our moral endeavor. How is the innate corruption of the will possible given Kant’s elaborate defense of the categorical imperative as the law of autonomy for rational but finite human beings? As a matter of fact, Kant offers his account of the corruption of the will in various places, but two accounts stand out prominently. First, in the *Groundwork*, Kant gives an account of our moral failure in terms of the subreption of the two types of maxims under what he calls “natural dialectic.” For example, toward the end of the first section of the *Groundwork*, Kant remarks that there is

a propensity to rationalize against those strict laws of duty and to cast doubt upon their validity, or at least upon their purity and strictness, and where possible, to make them more into accord with our wishes and inclinations, i.e., to corrupt their grounds and to deprive them of their entire dignity, something which not even common reason can in the end call good (G 4:406).

Kant then goes on to dub this propensity (*Hang*) a “natural dialectic,” a phrase that is surprising to the reader of the transcendental dialectic of the *First Critique*.

Kant further develops this story into a full-blown doctrine of radical evil in the *Religion*. In the first two books of the *Religion*, Kant famously claims that we human beings are morally corrupt, and goes on to locate the root of this corruption in what he calls the “propensity to evil.” He claims that this is due to our illicit use of our power of choice (*Willkür*). We become radically evil when we subordinate the moral law to our own self-interest. While we are no devils, totally denying the authority of the moral law, what we do when our will is corrupt is to invert the proper relationship between the incentives of the moral law and self-love. In other words, moral evil ensues when we incorporate both incentives into the maxim of our will, but make compliance with the law conditional upon self-love. Thus, for Kant the best epithet for human evil is not devilishness but perversity (*Verkehrtheit*) (R 6:36).

In this chapter, I will discuss Kant’s account of the will in his discussion of our moral failure as revealed in the natural dialectic of the *Groundwork* and radical evil in the *Religion*, respectively. I will suggest that Kant’s understanding of the way our will fails to be moral in the natural dialectic may be a precursor to a more sophisticated account of will as this is developed in his doctrine of radical

evil in the *Religion* but not quite “equivalent” to the latter. In Section 1, I will introduce the doctrine of natural dialectic in the *Groundwork* and how the Kantian conception of will together with its maxim is deployed in the process of a natural dialectic. In Section 2, I will discuss in detail how Kant develops a sophisticated conception of will in his elaboration of radical evil in the *Religion*. Section 3 then suggests that, while the *Groundwork* account of the corruption of the will proceeds by means of the fallacy of subreption predominantly committed at the level of individuals, in the *Religion* Kant gives an account of the corruption of the will by means of the two-tier structure of will, that is, the distinction between the *Wille* and *Willkür* at the level of the human species as a whole. Section 4 attempts to shed light on the exact relation between the two accounts of the will. In particular, I will argue that, *pace* Henry Allison, they are not “virtually equivalent.” Section 5 is a brief conclusion.

2 The Natural Dialectic in Common Human Reason

In order to understand the natural dialectic in G I, we need to consider the context in which it is introduced and discussed. It is well known that the first section makes a transition from common rational to philosophic moral cognition in identifying the fundamental principle of morality from common rational knowledge (G 4:392). What is peculiar in Kant’s discussion here is that he shows profound respect for the claims of common rational cognition about morality.² Common sense is not in need of sophisticated knowledge to distinguish what is right or what is wrong. People on the street have at least an equal chance of hitting the target as moral experts as far as ethical matters are concerned. “There is,” as Kant puts it, “no need of science or philosophy in order to know what one has to do to be honest and good and even wise and virtuous” (G 4:404).

But this does not mean that Kant was always in favor of the “plain common sense (*gemeinen Menschenverstand*)” in his philosophy. Kant simply could not accept such common sense as the irreversible and irrevocable touchstone of general truths. Reason at the level of common knowledge then is not without its own problem. Even though common sense is sound in principle in moral matters, it is

² This admiration for common moral reason is quite surprising to the reader of the *Prolegomena* because, in the preface to the latter, he launches a most incisive attack on what he calls “plain common sense (*gemeinen Menschenverstand*)” and the appeal to it by the so-called “common sense” philosophers such as Thomas Reid (1710–96) and J. Priestly (1733–1804).

also fragile and easily corruptible, i.e., weak and unreliable. Common sense “does not think its principle so abstractly in a universal form” (G 4:403). It simply does not show any understanding of its supreme principle of morality before its eyes when evaluating the moral value of actions. This is why it can easily lose sight of the principle. Kant is thus definite and quite emphatic about the limits of common knowledge of morality. It is innocent for sure, but, for the very same reason, it easily gets astray and confused.

Most importantly for our purpose here, there is always a possibility of committing the fallacy of subreption—the error that consists in substituting different kinds of terms or concepts—in common sense, for the latter easily confuses what one does with what one feels. Kant points out that this is “an illusion that even the most practiced cannot altogether avoid” (CPrR 5:116 cf. G 4:451). In our less guarded moments, “we easily look upon what we ourselves do as something that we merely passively feel and take the moral incentive for a sensible impulse” (CPrR 5:117). Common moral cognition does not know itself—in fact, it does not look upon itself—and, as a result of this lack of self-reflection, becomes an easy prey for naturalistic pitfalls in ethics.³ In this respect, we may even say that it is inherently incomplete. Its cognition is piecemeal and fragmentary. Accordingly, it does not know how its various elements are related to each other.

It then appears that the problem with common sense is that it simply does not have the supreme principle of morality before its eyes when appraising the moral value of actions. Accordingly, common sense can easily lose sight of it. This is fundamentally due to the fact that we are finite sensible beings. As products of nature, we all desire happiness, but in many cases our desire for happiness gets into conflict with the strict demands of reason. And it is usually the demands of reason, not the desire, that are sacrificed. Furthermore, our various needs and inclinations that we have as natural beings oppose the precepts of reason in all kinds of ways. The opposition between what is rational and what is sensible is in a way inevitable for limited beings like us.

As a result, common reason inevitably leads to an illusion of practical cognition which Kant introduces and discusses under the rubric “natural dialectic” (G 4:405. Cf. CPR, “Natural Dialectic of Human Reason,” A609/B697-A704/B732). Common human reason is constantly bombarded with the unruly and exorbitant demands of desires and inclinations, which are the products of experience, i.e., the brute nature in us. The fundamental precepts of common human reason thus get easily spoiled by such demands. Indeed, the dictates of reason in the form of duties are always bound to be counterweighed by the needs and inclinations,

³ For more on this, cf. Kim 2016, xxii-xxvii.

which are the undeniable and unavoidable conditions of human existence. This physico-psychological condition then gives rise to the natural dialectic, which is “a propensity to rationalize against those strict laws of duty and to cast doubt upon their validity, or at least upon their purity and strictness, and where possible, to make them more into accord with our wishes and inclinations, i.e., to corrupt their grounds and to deprive them of their entire dignity, something which not even common reason can in the end call good” (G 4:406).

At first blush, while the passage seems to indicate the presence of a persistent resistance our will shows to the demands of morality, it also quickly gives rise to some serious puzzles. Why does Kant call such a propensity a dialectic? According to Kant in the first *Critique*, a dialectic is concerned with a fallacious reasoning on the part of human reason that generates a transcendental illusion (A297/B354). Then the immediate puzzle for us is: what is the role that will plays in the way that this propensity generates a fallacious reasoning?

There is no question that what Kant calls “natural dialectic” shows that common reason easily sides with the brute forces of desires and inclinations, and gets comfortably settled in the easygoing life of the senses at the price of the ‘natural end’ of man, i.e., morality (G 4:396–7, CJ 5:442ff). Note that, even though this dialectic would not exist without our brute, sensible nature, it is not exactly due to the latter. What is responsible for the occurrence of this dialectic is none other than common reason itself, for it is common reason—i.e., our own reason at the common, ordinary level of consciousness—that sides with the prodding of the sensible nature in us, thereby generating the dialectic. Accordingly, it is the work of reason in action—i.e. the will—that is in conflict with itself. Thus, the epithet “dialectic.”

In the “Introduction to the Transcendental Dialectic” of the first *Critique* Kant explains how the metaphysics give rise to the flight of fancy quite illegitimately. Whereas the Transcendental Aesthetic focuses on the workings of sensibility, and the Transcendental Analytic on those of the understanding, the Dialectic here is devoted to an examination of the faculty of reason. In particular, Kant is concerned to show that the use of reason, when carelessly employed, could easily pave the way for drawing erroneous conclusions in metaphysics. Reason, which Kant describes as a capacity for drawing inferences, is here illicitly employed when its principles are transcendent (i.e., when they take away the limits of experience in a bid to overstep them). Transcendental Dialectic then uncovers the errors of metaphysics which Kant calls the “transcendental illusions” (ibid.). These illusions are generated when reason makes judgments on its own, using its own pure Ideas, not concepts of the understanding. The Ideas here include those of God, freedom, immortality and the unconditioned world, the world seen from no point of view.

These illusions, however, are quite “natural” and “unavoidable” because they are grounded in the “very nature of human reason” (A298/B354). This simply reflects a tendency in us to “take a subjective necessity of a connection of our concepts [...] for an objective necessity in the determination of things in themselves” (A297/B354). Indeed, Dialectic is defined as “the logic of illusion [Schein]” (A293/B350), as opposed to the Analytic, which is the logic of truth.

How do you resolve this dialectic? Kant apparently suggests that *only* by critically examining the power of practical reason can philosophy appropriate commonsense knowledge of morality and present it more completely and systematically. Such a complete “critique of reason” can provide recipes for a satisfactory resolution of such natural dialectic. Only transcendental deduction of the moral law can prevent the natural dialectic from occurring, dispelling all its doubts and prejudices. Thus, Kant devotes the entire Section III of the *Groundwork* to such an a priori proof of the moral nature in us by recourse to freedom found in and exercised by rational beings.⁴

3 Radical Evil

Now there is little room for doubt that this natural dialectic of common reason resonates well with the doctrine of radical evil in human nature that Kant later develops in the *Religion within the Limits of Reason Alone*. As social beings, we are acutely aware of the weight of morality, i.e., we are all conscious of the binding force that the moral law exerts on all of us including the most heinous scoundrel (G 4:394). To be sure, there is no question Kant firmly believes we are morally good when we adopt the moral law as our maxim and proceed to act on it. However, we often find ourselves rebel against the strict demands that morality makes upon us. In other words, we are constantly threatened by the temptations in our natural environments and easily coaxed into committing moral evil. As a result, we easily fail to adopt the moral law as the supreme maxim and instead end up taking incentives from the senses into our maxim. We can be morally evil in this sense. As Kant himself admits, this moral evil is not, however, the direct work of our biological nature. Even though it may be ubiquitously found in all humans, this clearly is not an intrinsic part of our essential nature. In other words, for Kant, moral evil is not essential to human beings, although it is universally found in all human beings, “even the best (R 6:30).”

⁴ For more on this, cf. Kim 2016, ch. 5.

What then is the source of moral evil in us? Clearly in the *Critique of Practical Reason*, Kant observes that human beings are “finite rational beings” (CPrR 5:23). This implies we are always subject to inclinations that often defy the moral law and can experience an inner conflict in our practical life. After all, the moral law is always perceived as a *categorical imperative* for us humans. As Kant puts it,

[f]or, being a creature and thus always dependent with regard to what he requires for complete satisfaction with his condition, he can never be altogether free from desires and inclinations which, because they rest on physical causes, do not of themselves accord with the moral law, which has quite different sources; and consequently, with reference to those desires, it is always necessary for him to base the disposition of his maxims on moral necessitation, not on ready fidelity but on respect, which demands compliance with the law, even though this is done reluctantly. (CPrR 5:71)

As finite beings, humans are driven and conditioned by their basic bodily desires. Yet, as rational beings at the same time, they cannot turn a deaf ear to the “voice of reason” and ignore the commands of the moral law (CPrR 5:32). This immediately suggests that there are two “direct opposite” principles in us that can be adopted as the “determining ground of the will”: one is “the principle of morality” and the other “the “principle of self-love” (CPrR 5:26, 32). Now, “the principle of self-love” or happiness can indeed furnish maxims, but it can never serve as “laws of the will,” because it is incapable of giving rise to “practical laws” that are objective, universal practical rules, and which are the determining ground of “the will of every rational being” (CPrR 5:21, 22, 33). Thus, maxims furnished by the principle of self-love are always contingent and necessarily have a limited validity.

The pivotal question that we must face now is: how can a human being adopt a practical law, which is objective and universal, as his or her maxim, which is only a subjective principle? Kant in fact does not hesitate to answer: “If a rational being is to think of his maxims as practical universal laws, he can think of them only as principles that contain the determining ground of the will, not by their matter but only by their form” (CPrR 5:24). In other words, only “the [universal] form of the maxim” can be thought as a practical law. Forming the determining ground of the will, the practical law in effect tells us to “act only in accordance with that maxim through which you can at the same time will that it should become a universal law” (G 4:431). If a person adopts this law as his or her maxim, then this person is morally good.

However, as pointed out above, to make the moral law as the “supreme condition” of the principle of self-love is not an easy task for human beings. They are easily tempted by what “promises agreeableness” (G 4:432). Kant makes the fol-

lowing observation concerning the struggle of the finite rational beings in the *Religion*:

Man (even the most wicked) does not, under any maxim whatsoever, repudiate the moral law in the manner of a rebel (renouncing obedience to it). The law, rather, forces itself upon him irresistibly by virtue of his moral predisposition; and were there no other incentive working in opposition, he would adopt the law into his supreme maxim as the sufficient determining ground of his will; that is, he would be morally good. But by virtue of an equally innocent natural predisposition he depends upon the incentives of his sensuous nature and adopts them also (in accordance with the subjective principle of self-love) into his maxim. (R 6:36)

This passage confirms that, as finite rational beings, we constantly struggle to follow the moral law. Further, Kant here suggests that incentives provide “the material of the maxims” (R 6:31). Both the moral law and “the sensual impulse” are incentives competing to determine the will. Unfortunately, it is simply not possible for us to take up both the incentive of the moral law and also the incentive of the sensuous nature in our supreme maxim when we want to be moral. For human beings cannot “be morally good in some ways and at the same time morally evil in others” (R 6:20). In other words, they cannot be partially good and partially evil, or good and evil at the same time (R 6:31). There is simply nothing intermediate between a good will and an evil one.

Now, if a person adopts the moral law into his or her supreme maxim, then that person would be morally good. However, if any incentive deviating from the moral law is adopted into the supreme maxim of a person, then he or she would be morally evil (R 6:20, 31). Such evil is radical, Kant asserts, because “it corrupts the ground of all maxims” (R 6:32). This is why Kant is naturally led to answer his transcendental question concerning our innate corruption of the heart by inquiring into “the *origin* of moral evil” (R 6:35; my emphasis). For instance, when a person commits evil acts, we are able to infer a priori “an underlying evil maxim” in him. Further, from this evil maxim, we can infer the presence in the agent of an underlying common ground, itself a maxim, of all particular morally evil maxims” (R 6:16). This “ultimate subjective ground of the adoption of maxims” is called “disposition (*Gesinnung*),” which is a maxim itself (R 6:20). Kant even goes so far to suggest that this disposition is “innate” because “it is posited as the ground antecedent to the very use of freedom in experience [...] and is thus conceived of as present in man at birth—though birth need not be the cause of it” (R 6:17). The bottom line is that no matter how hard we may try, we always feel some sort of resistance to the demands of morality. Such a tendency always lurks within us as a potential source of corruption.

Furthermore, we can see that, for Kant, the relevant disposition is a timeless ground of a person's maxims at any given point of time as it must work "antecedent" to our exercise of freedom in the course of experience. Kant, however, also insists that the disposition in question is nothing but our character that is based on the "subjective ground of the exercise (under objective moral laws) of man's freedom in general" (R 6:20–21). Kant never tires of voicing his conviction that we are responsible for our own disposition or character (whether good or evil), even if its fundamental ground is innate (R 6:21). Evil is then radical in the sense that our responsibility for it originates from the rational choice of the underlying ground on the part of our faculty of volition.

It is important to note that, for Kant, the faculty of volition or desire (*Wille*) has two different senses, a broad sense and a narrow sense. In the narrow sense (as the *Wille* proper), it refers to the practical power that legislates laws as the "faculty of desire whose inner determining ground, hence even what pleases it, lies within the subject's [practical] reason (MM, 6:213)." *Wille* then provides the ground that can help determine the choice of action (MM, 6:213), and, through it, an agent can formulate categorical imperatives. *Wille* stands in contrast with the elective will (*Willkür*), which is the free power of choice (together with which it forms the will in the broad sense) to choose, decide, wish, and formulate maxims presented to it by the *Wille* as imperatives. Hence, whether or not an agent is wholly good or evil is determined entirely by "a free power of choice (*Willkür*) and this power [...] on the basis of its maxims [which] must reside in the subjective ground of the possibility of the deviation of the maxims from the moral law" (R 6:29).

Thus, either the incentive of the moral law or the incentive of egoistic self-conceit is sufficient for the agent to be either morally good or morally evil. When the propensity to subordinate the moral law to the governing maxim of self-conceit is taken up within the mindset or disposition (*Gesinnung*) as a governing maxim, the agent's character as a whole is corrupted and evil.

In the *Religion*, Kant describes our choice of the supreme maxim in terms of "subordination" between the two incentives, i.e., the moral law or the principle of self-love. Only one can be "*the condition of the other*" (R 6:36). We are morally good when we subordinate the principle of self-love to the moral law. This then represents the "moral order of the incentives." On the other hand, we are morally evil when we adopt the two incentives in the opposite direction (R 6:31). To make the moral law the "supreme condition" of the principle of self-love is not an easy task for human beings.⁵ As Kant puts it,

5 On this evil tendency in us Kant also suggests that "to think of oneself as a freely acting being

[c]onsequently man (even the best) is evil only in that he reverses the moral order of the incentives when he adopts them into his maxim. [...] he makes the incentive of self-love and its inclinations the condition of obedience to the moral law (R 6:36).

In other words, moral evil is what happens when we reverse the subordination of the incentive of self-love to the moral law. Kant then suggests that this inversion of the moral order of these two incentives is rooted in our native propensity. As Kant suggests, we have a propensity which “does lie in human nature,” i.e., “there is in man a natural propensity to evil (*Hang zum Böse*) (R 6:29, 32).”⁶

We can thus see Kant give an account of the nature of the propensity (*Hang*) to evil as a disordering of incentives. As opposed to other vices, this propensity is essentially perversity, and stands in contrast to frailty (*fragilitas*) and moral impurity (*impuritas*, *improbilas*). Perversity (*perversitas*), unlike frailty, is not a mere weakness and an inability to resist sensuous inclination (R 6:29). And unlike impurity, it is more than merely obeying the moral law from alternative motivations (instead of a sense of duty). Instead, perversity must be understood as the reversal of “the moral order as regards the incentives of a free power of choice” (R 6:30). The propensity to evil becomes manifest when human beings, with their elective power (*Willkür*), choose to act in accordance with the incentive of self-conceit, which stands in opposition to the incentive of the moral law (R 6:36).

Note that the radical evil in us cannot be made fully intelligible (cf. Wood 2010, 147). Each of us has an overriding reason to act against it yet we end up being evil by way of giving higher priority to the incentives of self-love than to that of the moral law. Kant thus admits that “the ultimate subjective ground of the adoption of moral maxims is inscrutable” (R 6:21). In other words, we do not know what the ultimate ground is for a person to adopt the reversed moral order. It is a sheer mystery that we fall into radical evil on our own. Radical evil is thus inscrutable in this sense. Further, as we observed earlier, Kant

yet as exempted from the one law commensurate to such a being (the moral law) would amount to the thought of a cause operating without any law at all” and that is a contradiction (R 6:36), he adds.

6 Generally speaking, a propensity (*Hang*) is an innate yet non-necessary feature of every person that serves as a motivation for action in distinctively human affairs. However, unlike a basic predisposition (*Anlage*) (e.g. humanity, animality, and personality), a propensity can be represented as having been acquired by habit if it is good, or, if it is evil, as having been self-inflicted (R 6:29). It demonstrates a tendency to respond or act in a particular manner, either in accordance with, or in tension with the moral law. Taken together, both predispositions and propensities serve to form an individual’s mindset or disposition (*Gesinnung*), for the full development of which each human being is responsible.

always insists that, although radical evil is an inextirpable feature of human beings (because it is universally found), we are morally responsible for it. In other words, he claims that the fundamental disposition that gives rise to radical evil is still of our own making. We have a natural propensity to moral evil yet this is attributable to us because of our disposition, which is of our own making. Moral evil is then something we freely bring upon ourselves by the exercise of our power of free choice (*Willkür*). Moral corruption on our part then means the corruption of our *Willkür*. For Kant, then, “the source of evil [...] can lie only in a rule made by the *Willkür* for the use of its freedom, that is, in a maxim” (R 6:17). This is why Kant attributes radical evil to our agency. Further, this also paves the way for Kant’s cautious hope that, while this volitional corruption is a universal feature of human nature, it can be clearly countered by “the command that we ought to become better human beings” (R 6:45), which originates in the *Wille*.

At the same time, Kant also suggests that the fundamental disposition (*Gesinnung*) that serves as the ultimate subjective ground of the adoption of maxims is “acquired” because it is also “adopted by the power of free choice (*Willkür*)” but not “in time” (R 6:25). As he puts it:

Since, therefore, we are unable to derive this disposition, or rather its ultimate ground, from any original act of the will in time, we call it a property of the will which belongs to it by nature (although actually the disposition is grounded in freedom) (*ibid.*).

In other words, our *Willkür* can freely choose to acquire the fundamental maxim responsible for the selection of all our maxims. Since it does not choose in time, it then appears that our will chooses its fundamental maxim “innately.” When the will adopts the principle of self-love into the fundamental maxim, then “the ground of all maxims,” that is, the disposition, is corrupted, because the moral order of incentives is reversed (R 6:36). In this respect, we are naturally disposed to subordinate the incentive of the moral law to the incentive of self-love. Our fundamental disposition thus leads us to our inextirpably evil character in this sense.

It must be pointed out, however, that, despite the inextirpable, inscrutable nature of radical evil in us, Kant does not seem to despair of our human condition so severely. As he puts it, “the human being (even the worst) does not repudiate the moral law, whatever his maxims, in rebellious attitude (by revoking obedience to it)” (R 6:36). In other words, when we subject our self-love to reason, we do not have a positive incentive to violate the law; we do not seek evil for the sake of evil. We are no sheer devils. We do not reject the moral law simply because it is the moral law (cf. Michaelson 1990, 73). Better yet, we are able to recover from this radical evil. We are still morally capable of reversing the wrong-

headed order among our incentives. There is a “seed of goodness” remaining in the most wicked scoundrel (R 6:20). Kant thus points out that in spite of our evil heart, we not only have the native predispositions (*Anlagen*) to animality but also the predisposition to humanity as well as personality. We are thus natively in possession of a germ of goodness (R 6:45). We become morally corrupt when we invoke the free act of *Willkür* and invert the proper relationship between the moral law and the self-love, ending up making the compliance with the moral law conditional upon self-love. Nevertheless, it remains true that we are still aware of the binding force of reason. Our moral corruption cannot remove our *Wille*. Even the most heinous scoundrel is aware of the force of the moral law. After all, the moral law commands that we ought to be a better men and ‘ought’ implies ‘can’ (cf. Palmquist 2000, 159).

4 The Corruption of the Will in Natural Dialectic and Radical Evil

It is then beyond any reasonable doubt that, for Kant, our sensibility itself is not the direct source of moral evil. We cannot blame our biological nature for being a voluntary accomplice to the crime of moral corruptions. It follows that the natural dialect does not arise because the double nature is simply present in us, i.e., the sensible and rational natures in humans. Indeed, Kant is never an enemy of sensibility. It is not so much a direct result of the double nature in human beings as the consequence of reason’s voluntary acquiescence of the interests of the senses. In other words, moral evil rather comes from our own free decision to incorporate the incentives of sensibility into our maxim. This is why the natural dialectic does not give rise to any straightforward antinomy between the two heteronomous sources, i.e., the claims of duty and the claims of inclination. The natural dialectic rather arises as part of our natural skepticism about the validity of the moral law. Why should we be moral, we ask? In other words, why must the universal validity of our maxim as a law be the limiting condition of our action (G 4:449)? In this context, there is no question then that the natural dialectic stems from our common but practical reason. For empirically conditioned practical reason naturally and unavoidably challenges the claims of pure practical reason, finally overreaching itself and usurping the proper place of pure practical reason. This suggests that the natural dialectic points to a *conflict of reason with itself* because it shows the opposition of empirical practical reason with pure practical reason.

How then do we go about resolving this conflict? In G III, Kant suggests that we have to give a proper justification of the fundamental principle of morality, i.e., only the authentication of the categorical imperative by way of a transcendental deduction of freedom and moral law will do the trick. But this can be achieved only when we make a transition to a critique of pure practical reason from common moral understanding via metaphysics of morals (G 4:446–458). The latter is the answer to the question of the justification of morality, which is also the answer to the natural dialectic.

In the *Groundwork*, the natural dialectic is characterized as the source of corruption of the will, and here corruption has something to do with its *motivational* ground. In other words, the natural dialectic stems from our skepticism, on the part of each rational agent, about the reason as to why each of us has to accept and follow the claims of pure practical reason. The dialectic thus deprives pure reason of its *purity*, leaving it impotent.⁷ Accordingly, Kant's point suggests that each of us has a native tendency to challenge the motive of duty. Each of us naturally imagines that pure reason is not practical itself. This is of course illusory. We can also see that the natural dialectic predominantly takes place at the level of individual human agents.

In the *Religion*, however, Kant proposes an account of radical evil at the level of the human species as a whole. When Kant employs the term “human being (der Mensch),” he does not mean a single individual but the entire race there. He asserts that human beings ubiquitously manifest radical evil that accounts for how we might freely decide to subordinate the basic commitments of freedom to self-love. Kant claims that, although radical evil is not essential to human beings, e.g., it is not analytically contained in the definition of human beings, it is nevertheless necessarily present in all human beings, “even the best.” Such an evil is supposedly inextirpable because it goes beyond the “human forces (R 6:37).” We are existentially constituted to express an ineluctable resistance to the demands of morality and thus we may say the propensity to evil always lurks within us as a potential source of corruption. Nevertheless, Kant adamantly insists that we are morally responsible for it, i.e., it is something we freely bring upon ourselves and continuously embrace. How is this possible?

In order to answer this question, we have to begin with the observation that, for Kant, humanity as a whole has a specific moral nature. In particular, we always feel some sort of resistance to morality as part of our nature. But this resistance is not due to our biological nature. It is not due to our nature as animal

7 This is thus the opposite of impurity, which constitutes a trio of moral evil together with frailty and perversity. Impurity is an evil that consists in the corruption of our humanity (R 6:30).

beings. Our predisposition to animality is not to blame. As rational beings with needs, however, we necessarily consider and make every effort at procuring and promoting our material conditions vis a vis others (cf. Wood 2010, 158). In other words, our predisposition to humanity plays a direct role in contributing to a voluntary incorporation of the incentives of self-love into our supreme maxim. This is most dramatically presented in our social life where we have to compete with each other to promote our own happiness, even though we also need each other and form a society. Our social character then has hands on our propensity to evil. In his historical speculations,⁸ Kant interprets the corruption of the will at the collective level on the part of human species in terms of a naturalistic, Rousseauian teleology. Radical evil then may be cast in terms of what Kant identifies as “unsociable sociability” (*ungesellige Geselligkeit*, see 8:20). Nature has supposedly imbued human beings with a kind of “unsociable sociability” that engenders a continuous social struggle for the esteem of others. It arises within the human agent from interactions within society. Because of our predisposition to humanity, we possess a natural tendency not only to compare ourselves with others, but to compete with each other as a means of deriving our own self-worth. From our social interactions, we learn to give preference to our own concerns and needs, or self-conceit (R 6:26–27).

This unsociable sociability becomes manifest in our tendency to exempt ourselves from the moral law while expecting others to follow it, treating others as means to our ends rather than as ends in themselves. Accordingly, in human competitiveness, we seek to compare and gain mastery over others, making our own preferences the basis for our governing maxim. Hence, by virtue of living in a community and in our need for sociality, the shortcomings of our basic predisposition to humanity account for our self-conceit. Our social interactions end up serving as a kind of fertile breeding ground for radical evil. Thus, on this picture, the propensity to evil is simply part of our nature as social beings and is aggravated by our proximity to each other, the existence of which is evident from an observation of unsociable sociability when it occurs in human society. It is a universal feature shared by every human being, yet it does not require holding that each individual necessarily possesses this feature. Kant then concludes that radical evil is something that the human species as a whole freely elects to embrace, and for which we bear collective responsibility because we inculcate in ourselves a natural propensity to make exceptions to the demand of the moral law in circumstances when such an exception seems

⁸ For example, *Idea for a Universal History from a Cosmopolitan Point of View, Conjectural Beginning of Human History, and Perpetual Peace*, et al.

to be in our own favor and contribute to our well-being vis à vis others in the society (cf. Sussman 2010).

At the more metaphysical level, however, Kant points out that human nature straddles the divide between the pure (or a priori) and the empirical (a posteriori). Because of our membership in the intelligible order, there is always hope for human beings to improve themselves. Echoing his observation of the Bible, Kant suggests:

The absolutely first beginning of all evil represented as inconceivable by us [...] but man is represented as having fallen into evil only through seduction, and hence as being not fundamentally corrupt [...] but rather as still capable of an improvement (R 6:44).

This fall from the state of innocence into that of evil, fortunately, is never a hopeless condition. We cannot completely eradicate perversity in us, but that does not mean we should altogether abandon all our hope for moral improvement. As he puts it,

for man, therefore, who despite a corrupted heart yet possesses a good will, there remains hope of a return to the good from which he has strayed (R 6:39).

In other words, we need not despair of the condition of humanity. To become morally better, a person needs to have “a revolution in his cast of mind” as well as “a gradual reformation in the mode of senses” (R 6:43). That is to say, this transformation is both intelligible and sensible: “For, the will stands between its a priori principle, which is formal, and its a posterior incentive, which is material, as at a crossroads” (G 4:413). Through this continuous and infinite progress, hopefully human beings can all strive to become morally better persons.⁹

The metaphysical basis, on the Kantian part, of the corruption of the will at the collective level is then that the subordination of the moral law to the incentive of self-conceit is an entirely timeless and intelligible “deed (*That*)” (R 6:31). This wholly intelligible deed is so called because it does not take place at any one point in time, but it is nevertheless the deed out of which all subsequent evil deeds in time arise. It is, as Kant states, a “subjective determining ground

⁹ Kant even claims that this hope may include the hope for divine assistance in performing the revolution (R 6:171). But if Kant is understood as claiming that we require assistance from God in order to become morally good, this necessarily generates a contraction with his fundamental assumption that human beings are free and fully responsible for being good or evil. Andrew Chignell proposes that Kant might “regard it as metaphysically possible for us to receive divine assistance in the intelligible realm while doing something for which we are fully responsible.” Cf. Chignell 2013, 216.

of the power of choice (*Willkür*) that precedes every deed and is itself not yet a deed” (R 6:31). It is then a formal act whereby the supreme maxim is adopted by our will (*ibid.*) This intelligible act then produces a propensity to evil in all the members of the human race (cf. Palmquist 2000, 155). Human beings approach their empirical circumstances having always already chosen the supreme maxim by which they will act, and so subordinate the moral law to the incentive of self-conceit. This is why human beings are generally susceptible to natural inclinations that go against the dictates of the moral law. Rather than naturally possessing a propensity to follow the moral law, humans instead possess a propensity to follow their own self-serving inclinations. Since, as we saw earlier, human beings are wholly good or evil by virtue of whether or not they choose a morally worthy maxim or a selfish maxim at the top of their hierarchy of maxims, this propensity must be evil and imputable to human agency.

Accordingly, even though this propensity to evil is “innate,” it can also be considered as being “brought by the human being upon himself” (R 6:28). That is to say, our nature is imputable if it does not follow the moral law. This propensity to radical evil then is both innate and acquired (R 6:25, 29, 43). Kant thinks that “a propensity to evil can only attach to the moral capacity of the will. But nothing is morally (i.e., imputably) evil but that which is our own deed” (R 6:31). The puzzle about the propensity is that it is ambivalent, i.e., it is both an act (*That*) and not an act. Kant obviously believes that the term “act” can refer to the will’s “exercise of freedom” to adopt “the supreme maxim” (*ibid.*). Alternatively, it can also refer to the actions performed through “the exercise of freedom [...] in accordance with that maxim” (*ibid.*). The former is “intelligible action” while the latter is a “sensible action” (*ibid.*). Kant explains that the innate propensity to evil concerns an “intelligible action,” therefore not an act performed in time, because it is “a subjective determining ground of the will which precedes all acts and which, therefore, is itself not an act” (*ibid.*). In other words, it refers to the will’s choice of the principle of self-love as the supreme maxim, and therefore not an empirical action.¹⁰ But when “actions themselves (materially considered, i.e., as regards the objects of the power of choice) are performed in accordance with that [supreme] maxim,” then we may say that the propensity to evil concerns a “sensible action,” too, since it is “the formal ground of every deed contrary to law” and resists the law materially and is then called vice” (*ibid.*).¹¹ Accordingly, the natural propensity to evil as an intelligible action “cannot be extirpated” (6:37), Kant suggests. In other words, the

¹⁰ Kant calls this propensity to evil “the original sin (*peccatum originarium*)” (R 6:26).

¹¹ This is the “derivative sin (*peccatum derivatum*)” (R 6:26).

original sin is ineradicable. Nevertheless, the propensity to evil as an intelligible act points in the direction of free adoption of the supreme maxim on the part of the *Willkür*, thus paving the way for our responsibility but the same propensity expressing the set of our sensible choices of concrete maxims in accordance with the supreme maxim signals that it is fixed. As a member of the intelligible order, we are responsible for the evil, but as members of the sensible order, we are universally inclined towards vice.

5 Natural Dialectic vs Radical Evil

Now, Henry Allison suggests that the natural dialectic as a propensity to rationalize against the claims of duty shows a “virtual equivalence to [...] a propensity to evil” (Allison 2011, 143f.) Allison does not give a precise sense of “equivalence” (nor one of “virtual equivalence”), but from the scattered remarks in his *Kant’s Groundwork for the Metaphysics of Morals: A Commentary*, we can reconstruct its rough meaning: Both natural dialectic and radical evil are described as a propensity in the sense of a self-imposed tendency in an agent to subordinate moral considerations to self-love (cf. Allison 2011, 202, n.59); both stem from our sensuous nature as their necessary but not sufficient condition (cf. Allison 2011, 203); it is the work of the will that is required to generate both natural dialectic and radical evil (cf. Allison 2011, 291); finally, a good will is compatible with both, but both involve a wrongful use of common human reason, whose resolution lies elsewhere (e.g., in a ‘critique’ beyond the scope of common reason) (cf. Allison 2011, 142, 144).

Even though there is no question that the doctrine of the natural dialectic provides elementary contentual as well as historical background to the more mature and advanced doctrine of radical evil in the *Religion*, the claim of ‘virtual equivalence’ seems to be plainly far-fetched if the claim entails that the latter is a simple restatement and expansion of the former. First of all, in the *Groundwork*, Kant clearly believes that this dialectic can be overcome by a transition to a proper philosophy. But what does this mean? Kant suggests that the universal validity of the moral law must be justified to silence the strong skepticism about morality stirred by the claims of senses. In other words, the complete resolution of the natural dialectic requires a carefully executed proof of our moral nature under what Kant calls “transcendental deduction.” But this can be carried out only in a complete critique of our reason (G 4:405). Otherwise, even common reason can find itself tangled in an illusion of a dialectic.

On the other hand, the resolution of radical evil does not need any transition to philosophically regimented enterprises of the kind we see in the *Groundwork*

III. After suggesting that “we can spare ourselves the formal proof that there must be [...] a corrupt propensity [to evil] rooted in the human being, in view of the multitude of woeful examples that the experience of human *deeds* parades before us” (R 6:32–33), Kant insinuates that we can resolve the problem of radical evil by a recourse to a practically oriented reflection on human nature. In other words, Kant tells us that we can achieve the ‘change of heart,’ i.e., a revolution in the constitutions of the supreme maxim, only by way of the notion of the practical ideal of moral perfection. According to Kant, each of us has, as an archetype, an ideal of a humanity pleasing to God, which is present in our reason (R 6:74–75).¹² This is an archetype of a perfect person, Jesus, within us. Just as the rational origin of the propensity to evil is inscrutable, so is that of the archetype. It is only when we have faith in the practical validity of that idea lying in our reason we may hope to become acceptable to God (and so be saved) through a “*practical faith in this Son of God*” (R 6:61). It is practical faith, not the a priori transcendental proof, that can initiate a moral revolution by way of giving rise to moral maxims. Our faith in the archetype of perfection and its incorporation of it into a good maxim then presumably leads to a conversion or a change of heart on our part. We may then embark upon the road of endless progress towards holiness (R 6:46–47). This, however, cannot take a gradual, temporal process. Thus, the anticipated moral revolution requires a radical change in our intelligible character.

Also, as opposed to the strong individualistic undercurrent in the *Groundwork*, Kant, in the *Religion*, suggests that radical evil is something that the human species as a whole (R 6:25) freely elects, and for which we bear collective responsibility.¹³ There is no question Kant has become adamant about the human vulnerability in his discussion of radical evil in us but there is also the lingering hope that we can achieve moral progress by a revolution. Above all, a good will is compatible with a propensity to evil. Since we have the double nature that straddles the divide between the intelligible order and the sensible order, there is always hope for human beings to improve themselves at least in their intelligible character.

Furthermore, the vice involved in the natural dialectic is limited to the frailty and impurity of the will, but the propensity to evil pinpoints our perversity. Evil does not so much consist in the character of incentives themselves as our free adherence to a reversed order among the incentives. In other words, evil ensues

¹² This is Kant’s philosophical interpretation of the figure of Jesus in the Bible (R 6:67–78).

¹³ Palmquist argues that the generic term “man” refers to the ‘entire race,’ not to the ‘single individual’ and goes on to point out that this feature is prominent in books III and IV (more than books I and II). Cf. Palmquist 2000, 150.

when we freely prefer the incentive of self-love to the incentive of the moral law in the determination of our supreme maxim.

Accordingly, it is inevitable we cannot accept Allison's suggestion as is. The natural dialectic cannot be "virtually equivalent" to the propensity to evil. While there is no denying that the latter is clearly based on our tendency to go against the strict laws of duty, this constitutes only a necessary condition of the radical evil in us. The human propensity in question is not simply that we evade the claims of morality whenever possible. The natural dialectic is a mere "moral illusion," i.e., a fantasy of the reign of heteronomy exposed in the *Analytic of Practical Reason*, but the perversity of human nature exposed in radical evil is our inexorable "reality" (cf. Beck 1960, 241). We are aware of our capacity of autonomy but we find ourselves actually inclined to the way of heteronomy, so to speak. But it also involves our explicit preference we have for the claims of sensibility to reason. We are the willing accomplices to the claims of sensibility in our moral struggle, so to speak.

Note that the natural dialectic is clearly different in nature from the antinomy of pure practical reason Kant discusses in the second *Critique*. It is plain that the natural dialectic does not lead to any antinomies. It is rather a weak dialectic (Klemme 2010, 24). But it is still dialectical because it "falls into sheer incomprehensibilities and self-contradictions, at least into a chaos of uncertainty, obscurity, and instability" (G 4:404). It is beyond any doubt that the natural dialectic begins with our desires and inclinations, i.e., our sensible nature. But it does not end with them. Indeed, our sensible nature provides a fertile but only necessary ground for rebelling against the strict demands of reason. Kant clearly holds that this tendency is ultimately rooted in our pursuit of happiness. In the early portion of the *Analytic of Practical Reason* of the second *Critique* where he presents the fundamental tenets of his account of empirically practical reason, its first two theorems state that "all practical principles that presuppose an object of the faculty of desire as the determining ground of the will are empirical and can furnish no practical laws" (CPrR 5:21), and that "they all are grounded on the principle of self-love or one's own happiness" (CPrR 5:23). So, Kant clearly believes that all material practical principles are of one and the same kind and come under the general principle of self-love or one's own happiness. This has the consequence that "all material practical rules put the determining ground of the will in the lower faculty of desire, and were there no merely formal laws of the will sufficient to determine it, then neither could any higher faculty of desire be admitted (CPrR 5:23)." Kant then goes on to prove his next theorem, i.e., "if a rational being is to think of his maxims as practical universal laws, he can think of them only as principles that contain the determining ground of the will not by their matter but only by their form."

For Kant, then, our pursuit of happiness provides only a seed for perpetuating the natural dialectic. For the latter to take place, our own willing cooperation and its incorporation into our maxim is needed, no matter how incessantly and tenaciously our sensible nature beseeches and petitions.

There is simply no royal road to morality as it issues “its precepts unremittingly, without thereby promising anything to the inclinations, and so, as it were, with disregard and contempt for those claims” (G 4:405). Indeed, Kant does not relent in holding that the claims of the senses are so “impetuous and [...] equitable (and refuse to be neutralized by any command.)” (G 4:405) This is why we need to make a transition to a metaphysics of morals from common rational knowledge of morality. Ultimately, like all dialectical inferences, the natural dialectic is to be resolved with an appeal to transcendental idealism, in particular, the deduction of the moral law based on the distinction between things in themselves and appearances.

This is also the reason why the natural dialectic does not arise in the second *Critique*. Instead, Kant points out a new way of a dialectic in the concept of the highest good as the necessary object of our volition. There is a conflict between our pursuit of happiness and our concern with virtue. We thus find ourselves caught between a pursuit of happiness as the motive of virtuous action and virtue as the ground for my hope for happiness as its reward. But this is not quite radical because it can be resolved with the insertion of the distinction between the things in themselves and the appearances. But radical evil in our will can be resolved only with the help of our redemption as the whole species.¹⁴

6 Conclusion

The natural dialectic in the *Groundwork* is a tendency or a propensity to rationalize against the claims of duty. This is responsible for the illusion of the supremacy of heteronomy over autonomy. This can be resolved by a recourse to transcendental idealism, i.e., a recourse to the distinction between appearances and things in themselves. On the other hand, radical evil expresses our innate tendency to choose the principle of self-love over the principle of morality as our supreme maxim. As such, it is inextirpable. It is not clear how we can extricate ourselves from this deeply rooted propensity to evil, but Kant clearly puts his hope on a moral redemption, starting with the compatibility of our good will with the propensity to evil (R 6:29, 77) together with a practical faith in

¹⁴ R 6:95-124. I cannot discuss this fully but for more on this, cf. Palmquist 2000.

the archetype of humanity, eventually going so far to invoke a divine assistance at the level of the whole human species. Thus, Kant's complex story of our original innocence and innate evil necessarily leads to religion.

In this respect, the doctrine of natural dialectic in the *Groundwork* only serves as the precursor to Kant's mature theory of radical evil in the *Religion*. They are not equivalent. However, both doctrines, no matter what the exact status of their mutual relation may be, clearly indicate that, for Kant, achieving morality amidst our innate propensity to evil decisively involves the committed operations of the will, in particular, the exercise of our free choice, in the embodied yet rational life on our part.¹⁵

Bibliography and Abbreviations

- CPR: Critique of Pure Reason
 CPrR: Critique of Practical Reason
 CJ: Critique of the Power of Judgment
 G: Groundwork of the Metaphysics of Morals
 MM: Metaphysics of Morals
 R: Religion within the Boundaries of Mere Reason

- Allison, Henry (2011): *Kant's Groundwork for the Metaphysics of Morals: A Commentary*. Oxford: Oxford University Press.
- Anderson-Gold, Sharon/Muchnik, Pablo (Eds.) (2010): *Kant's Anatomy of Evil*. New York: Cambridge University Press.
- Beck, Lewis W. (1960): *A Commentary on Kant's Critique of Practical Reason*. Chicago: University of Chicago Press.
- Chignell, Andrew (2013): "Rational Hope, Moral Order, and the Revolution of the Will." In: Eric Watkins (Ed.): *Divine Order, Human Order, and the Order of Nature: Historical Perspectives*. Oxford: Oxford University Press.
- Fichte, Johann G. (2005): *The System of Ethics*. Trans. D. Breazeale and G. Zöller. Cambridge: Cambridge University Press.
- Kant, Immanuel (1900-): *Gesammelte Schriften*. Vol. 1–22. Ed. by the Preussischen Akademie der Wissenschaften, Vol. 23 by the Deutschen Akademie der Wissenschaften. Berlin/New York: de Gruyter.
- Kant, Immanuel (1991-): *The Cambridge Edition of the Works of Immanuel Kant*. New York: Cambridge University Press.

¹⁵ An earlier version of this paper was read at the annual Korean Kant Society meeting at Sogang University in 2017 and also at the conference on the Concept of Will in Classical German Philosophy in Munich at LMU Munich in 2017. I am indebted to S. Palmquist and Ji-Young Kang for their comments on an earlier draft of this paper. I also thank the editors of this volume, in particular, Manja Kisner, for valuable suggestions.

- Kant, Immanuel (1996): *Critique of Practical Reason*. Translated by Mary Gregor. In: Allen W. Wood (Ed.): *Practical Philosophy*. Cambridge and New York: Cambridge University Press.
- Kant, Immanuel (1996): *Groundwork of the Metaphysics of Morals*. In: Allen W. Wood (Ed.): *Practical Philosophy*. Cambridge and New York: Cambridge University Press.
- Kant, Immanuel (1996): *Metaphysics of Morals*. Translated by Mary Gregor. In: Allen W. Wood (Ed.): *Practical Philosophy*. Cambridge and New York: Cambridge University Press.
- Kant, Immanuel (1996): *Religion within the Boundaries of Mere Reason*. In: Allen Wood/George Di Giovanni (Eds.): *Religion and Rational Theology*. Cambridge and New York: Cambridge University Press.
- Kant, Immanuel (1998): *Critique of Pure Reason*. Translated by Paul Guyer and Allen W. Wood. Cambridge and New York: Cambridge University Press.
- Kant, Immanuel (2000): *Critique of the Power of Judgment*. Trans. Eric Matthews and Paul Guyer. Cambridge and New York: Cambridge University Press.
- Kim, Halla (2016): *Kant and the Foundations of Morality*. Maryland: Lexington Press.
- Klemme, Heiner (2010): "The Origin and Aim of Kant's Critique of Practical Reason." In: A. Reath/J. Timmermann (Eds.): *Kant's Critique of Practical Reason: A Critical Guide*. New York: Cambridge University Press, pp. 11–30.
- Michaelson, Gordon (1990): *Fallen Freedom*. New York: Cambridge University Press.
- Muchnik, Pablo (2010): *Kant's Theory of Evil*. Lanham: Lexington Books.
- Palmquist, Stephen (2000): *Kant's Critical Religion*. Aldershot: Ashgate.
- Schelling, F. W. J. (1887): *Of Human Freedom*. Trans. James Gutman. La Salle: Open Court.
- Sussman, David (2010): "Review of Muchnik, Pablo, *Kant's Theory of Evil*." In: *Notre Dame Review of Philosophy* 7.19.
- Wood, Allen W. (2010): "Kant and the Intelligibility of Evil." In: Sharon Anderson-God/Pablo Muchnik (Eds.): *Kant's Anatomy of Evil*. New York: Cambridge University Press, pp. 144–172.

Günter Zöller

Eleutheronomy: Will, Law and Liberty in Kant's Esoterically Political Philosophy

Abstract: This contribution explores the juridical formative background of Kant's ethics by placing Kant's legal and ethical thinking in the context of the modern normative discourse about law and liberty. Section 1 outlines the intrinsically republican profile of much of modern political philosophy. Section 2 features the terminological and conceptual distinction between freedom and liberty. Section 3 addresses the juridico-ethical analogy underlying the normative nature of Kant's practical philosophy. While the contribution's title, "eleutheronomy," comes directly from Kant, its subtitle attributes to Kant, not an esoteric political philosophy, but a practical philosophy, including an ethics, that is inwardly, "esoterically," shaped and animated by key concepts taken from (modern) political philosophy.¹

"Liberty! Freedom!"

Shakespeare, *Julius Caesar*, III, 1

1 The Discourse of Classical German Philosophy

Classical German philosophy in its narrower extension from Kant through Fichte and Schelling to Hegel and in its wider scope that includes Schopenhauer and Nietzsche as well as Feuerbach and Marx, while being bent on abstraction, analysis and argument in method and procedure, also is geared at willing, acting and doing in intent and application. The systematic thinking – a thinking in systems – so central to the critical philosophy, to German Idealism and to radical post-idealism is never an end in itself but the strategically sought basis, groundwork and framework for human intervention and immersion in the real world. Moreover, the world so envisioned and encountered is not to be experienced

¹ The paper forms part of an ongoing research project on classical German political philosophy in general and on the latter's republican commitments in particular. For further details, see the bibliography at the end of the contribution. Throughout the paper, references to Kant's work are by volume and page of the Academy Edition ("AA"), except for *the Critique of Pure Reason*, which is cited according to the original pagination of the first and second edition ("A," "B," respectively)..

as alien and antagonist but is to be made into, or recognized to being already (as in Hegel's case), a natural and cultural sphere that meets and matches essential human needs, desires and ends.

The chief concept underlying the primacy of the practical pursued in classical German thought is that of the will (*Wille*) or of willing (*Wollen*). From the "pure will" in Kant through the "eternal will" in Fichte and willing as "primary being" (*Ursein*) in Schelling to the "will to life" in Schopenhauer and the "will to power" in Nietzsche, the will furnishes the very principle of the world of human existence as reflectively refined to its innermost practical core. The philosophical focus on the will that is, if not definitional, then at least characteristic of classical German philosophy typically is conjoined with the normative notion of freedom to form the idealist and post-idealist core conception of the free will or the freedom of the will. The freedom attributed to the will in classical German philosophy involves a conception of human existence as exceeding the given and the natural in favor of the made – more precisely, the self-made or human-made – and the supernatural – the practical in general and the specifically moral or the normative in particular.

A foremost feature of the linkage between will and freedom established, in various ways, by classical German philosophy is the role of reason in the formation of the will and its freedom. From Kant outright identifying the will with practical reason to Schopenhauer explicitly severing the will from reason, reason figures as the matching or counter concept to the will's fundamental function for self and world alike. Moreover, from Kant through Fichte and Schelling to Schopenhauer and Nietzsche the will in its generic capacity as the practical faculty is subject to further differentiation and articulation. Most importantly, Kant comes to distinguish between elective will and legislative will, with the former involving arbitrary choice (*Willkür*) and the latter rational self-legislation or autonomy (*Wille*).

In subsequent developments Fichte differentiates between the finite and infinite will, Schelling between the human and the divine will, and Schopenhauer between the psychic and the cosmic will. In a similar vein, the conceptuality of freedom correlated with that of the will undergoes differentiation and modification during the development of classical German philosophy. In particular, Kant, Fichte, Schelling and Schopenhauer, each in their own way, distinguish between negative freedom or freedom-from and positive freedom or freedom-for. Further variety accrues to the twin accounts of freedom and the will over the course of classical German philosophy through the specification of the twin concepts by area or domain of pertinence and application. The spectrum of subtitles is especially noteworthy in the case of freedom. Thus Kant distinguishes between "outer" or legal and "inner" or ethical freedom, in addition to contrasting "tran-

scendental,” “practical” and “moral” freedom. Fichte in turns distinguishes between “formal” and “material” freedom. Schopenhauer separates the “freedom of willing” from the “freedom of acting.”

In a related vein, the complex constellation of will, freedom and reason peculiar to classical German philosophy is undergirded by the introduction of pre- or proto-rational features that either prepare or counteract the rational formation of the will in view of its freedom. The crucial concept operative here is that of the drive (*Trieb*), which originates in contemporary natural history (*Bildungstrieb*, “formative drive”), from where it is taken over by Kant. Following the lead of the influential post-Kantian, K. L. Reinhold, who reconstructs the Kantian opposition of inclinational and rational motivation as the arbitrarily free choice between the “selfish” and the “unselfish drive,” Fichte has the “ethical drive” result from the suitable integration of the “natural” and “pure drive,” with freedom fashioned as the power to consent or dissent from the naturally or rationally driven ends. Schelling attributes to all of nature a drive-like yearning, torn between the attraction to and the repulsion from the divine center of things. In Schopenhauer the will itself comes to be identified with the elementary, extra-rational, “blind” drive for life, while Nietzsche makes the will the multiform principle behind all reality, always aiming at subjection and domination, under the guise of the “will to power.”

The primacy of the practical, the centrality of the will and the focus on freedom collectively contribute to making classical German philosophy, on the whole, an eminently practical philosophy. To be sure, the practical philosophy aimed at by classical German philosophy is still philosophy – based on thinking, driven by analysis, proceeding by argument and aiming at a comprehensive account of both self and world. But as practical philosophy, classical German philosophy, while espousing a theoretical and even speculative stand, is geared toward practical matters – chiefly individual and social life – and, moreover, geared at elucidating the forms and norms that guide and govern life. The practical principles in question may be universal, formal and apodictic, as in Kant, or expressive of substantial commitments, as in Hegel, or indicative of a future yet to be brought about, as in Marx. Moreover, the various versions of classical German practical philosophy are united by a common concern with freedom in general and with social freedom or freedom under conditions of community in particular.

In addition to its general character as an eminently practical philosophy, classical German philosophy, in most, if not all, of its chief representatives, exhibits a particular bent toward elucidating and enhancing the political or civic life form. In its sustained focus on matters of right, law and ethics, classical German philosophy typically also includes the transition from those practical do-

mains and their principles to the norms and forms of politics and economics. In Kant, the political dimension comes to the fore in the “metaphysical first principles” of public law in his late *Metaphysics of Morals* as well as in earlier occasional and popular pieces in the philosophy of history and in political philosophy. In Fichte the treatment of the pure principles of civil and public law (“natural law”) extends into the nationalized economy of the “closed commercial state.” In Hegel’s *Philosophy of Right*, which includes law and ethics as well as civics and politics, the political domain is structured into the complementary social spheres of “civil society” and the “political state.” In the later Schelling the philosophical account of the state forms part of a philosophy of history focused on the creation of religious myths and truths. In Marx’ historical materialism politics is aligned with cultural practices of all kinds and emerges as the only seemingly independent superstructure of an economic basis shaped by technological advances and their civic repercussions under the guise of property relations.

But the political sphere and its distinct concepts and precepts form not only an important area of application and instantiation for classical German philosophical thought. Political notions also inform much of classical German practical philosophy in general by providing terms and ideas that originate in the political sphere and lend themselves to the transfer into different, though related areas. To be sure, the originally political vocabulary and conceptuality employed, by way of analogous extension, to extrapolitical issues and areas do not amount to an overt or intended politicization of other practical matters. Rather the political paradigm provides the theoretical tools for thinking about practical matters in general and moral matters in particular under the influence and inspiration of the philosophical discourse about politics.

The ancient precedent for the modern (*sit venia verbo*) politicomorphism, as evident in much of classical German philosophy, is Plato’s sustained analogy between the city and the human being in the *Republic*. Ostensibly employed as a means of elucidation, the configuration of ethical principles (“virtues”) in strict analogy to those of the city-state (*polis*) in its normative constitution (*politeia*) – with the comprehensive character of justice (*dikaioσύνη*) at the center of both – essentially reflects the formally and functionally identical set-up of civic and psychic life. Much of modern moral philosophy follows the Platonic methodological lead by introducing originally political concepts, distinctions and even doctrines into other, though related areas, foremost among them moral philosophy. The move is motivated by the Platonic insight, as conveyed through Socrates in the *Republic*, that the intricate inner life of a human being, given its small scale, is best studied and assessed by being construed on an analogy with the outward life of human beings in community with each other, profiting of the latter’s larger

scale. In modern philosophy, including the high modern development of classical German philosophy, the general epistemic enhancement of moral matters by appeal or recourse to political precedents is further motivated by the significant development that political life and philosophical reflection on it undergo in this period.

In its radical character and revolutionary impact modern political philosophy – from Bodin and Grotius through Hobbes and Locke to Montesquieu and Rousseau, to name only the most outstanding contributors before Kant and Hegel – is rivaled in its formative influence on modern philosophy in general only by modern science (“natural philosophy,” “natural history”) from Kepler and Galileo through Newton to Buffon and Lavoisier. The political realities that shape and challenge modern political thought reach from the Protestant reformation and the resultant confessional pluralism through the rise of the centrally governed territorial state and the associated consolidation of political power to the emergence of constitutional monarchy and the correlated conceptions of citizens’ and human rights and the capitalist economy as well as the political revolutions in North America and France.

While a major medium for responding to those political development is the law – Roman Law (*ius civile*) as well as natural law (*ius naturae*) – and its reflection in legal philosophy or philosophy of law, narrowly legal issues often assume wider significance through their political character and scope, especially in the sphere of public law (*ius publicum*) and its core concern with constitutional or state law and international or peoples’ law (*ius gentium*). Accordingly, even the lawyers among the modern political philosophers – including Bodin, Grotius and Montesquieu – employ legal issues and tools in the service of an essentially political conception of the nature and function of law in civil society.

To be sure, the very designation “political philosophy” is all but absent from the work of the early modern political philosophers and from that of their critical followers in classical German philosophy as well. In the absence of a narrowly disciplinary understanding of the term, the designation “political philosophy,” when applied to classical modern philosophy, including its culmination in classical German philosophy, takes on a wider meaning. In view of the generally practical and specifically political character of much of modern philosophical thought, the term refers not to a narrow field of philosophy but to an entire way of regarding and treating philosophy – as devoted to an assessment of life in the modern political community (polity), or “the state,” and concerned with freedom in general and freedom in the formation and exercise of the will on the part of the modern individual in particular.

2 Freedom and Liberty

The political origin of core concepts operative in modern moral philosophy, including classical German philosophy, is already apparent in the very vocabulary employed in much of modern thinking about customs, morals and ethics. In particular, the psychology, anthropology and sociology that informs modern moral philosophy typically draws on technical terms of a political nature: “power,” as in “power of the mind,” “law,” as in “moral law,” “rule,” as in “popular rule,” “obligation,” as in “moral obligation” or “will,” as in “moral will” – all derived from the language of politics, albeit adapted and assimilated to the peculiarities and requirements of the case at hand. But the political profile of classical German philosophy runs deeper yet. Not only do concepts from politics and political thought serve to illuminate moral matters in general by drawing comparisons, establishing analogies and furnishing metaphors. The very discourse about moral matters is animated and articulated by its political counterparts or rather role models. When classical German philosophy relies on political imagery for addressing and assessing other areas of practical philosophy, chiefly moral matters, the recourse is not auxiliary and optional but formative and even constitutive, if not essential.

The systematic reliance of modern practical philosophy in general and of classical German practical philosophy in particular on schemes and functions from political philosophy is especially striking in the case of the central concepts of modern political as well as moral thinking, viz., freedom and willing. To be sure, each of the two concepts typically carries multiple meanings that vary from one context to another. Semantically speaking, “political freedom” is entirely different from “moral freedom,” as is the “moral will” of an individual from the “political will” of a people or some part of it. Still, the different uses of the same term share more than a verbal shell. In a genealogical perspective, the seemingly different meanings of the same term can be traced to an original function of the term, prior to its diverging employment in different domains. Moreover, the original meaning may be closer to one such separate area than to another. In particular, the original usage may be political, or proto-political, in character and establish a political primacy of sorts in the chain of semantic expansion subsequently undergone by the term.

To be sure, the political origin of subsequently ethical terms, such as “freedom” and “will,” does not amount to a semantic reductionism that obliterates a subsequent change of signification and a concomitant separation of distinctly different areas of use. But the deeper affinity between different usages and their correlated meanings, together with the factual primacy of one domain –

or its predisjunctive precursor – over another, suggests a formative function of the earlier and more original domain for the very establishment and differentiation of a subsequent one. In particular, upon closer examination, key terms such as “freedom” and “will,” as employed in classical German philosophy, may exhibit a depth semantics that belies manifest differences and reveals a dynamics of differentiation based on a prototype underlying ostensible opposition. Drawing on a term employed by Kant to designate the inner workings of organic beings and subsequently used by W. v. Humboldt to capture the internal functioning structure of language, the para-political deep structure of moral notions such as freedom and will could be called their “inner form” (*innere Form*, AA 05: 376).

A methodological precedent for the proposed exploration of the predominantly political origin of non-political concepts is Nietzsche’s mature project of a “genealogy of morals” (*Genealogie der Moral*). Based on his training in classical philology and combining etymological and ideological critique, Nietzsche seeks to explore the extramoral past of presently moral concepts. Nietzsche’s focus though is on the alleged “reversal” of value systems and employs the genealogical approach to the establishment and exchange of moral value systems as a case study in the universal reign of the “will to power” as a force that subjects phenomena of all kinds to a sequence of domineering interpretations and reinterpretations. Moreover, as a project in the tradition of “radical Enlightenment” (J. Israel), Nietzsche’s genealogy of morals serves the polemical purpose of unmasking and denouncing seemingly objective modes of evaluation by exposing their underlying power politics – a project that was to find its continuation and extension in the later work of M. Foucault. By contrast, the proposed perspective on the political behind the practical in classical German philosophy involves functional analysis rather than ideological critique and aims at uncovering the political deep structure that informs and sustains the very functionality and integrity of classical German philosophy as a paradigm of modern philosophy in general and of modern political philosophy in particular.

The political profile behind the conception of freedom in modern philosophy is best explored by drawing a distinction that is linguistically available in English but that has its functional equivalents in other languages as well. Due to England’s past as an island sequentially invaded by a number of forces and peoples and occupied or settled by several of them, most notably the Anglo-Saxons and the Normans, the English language is especially rich in vocabulary, as attested by the extent of entries in the standard reference work on the language, the *Oxford English Dictionary* (OED). In particular, the earlier import of the Saxons’ Germanic language and the later overlay of the Normans’ French language have introduced a dual lexicon into the English language that has a distinctly socio-political character. Typically, the concrete, manifestly material meaning of a con-

cept or notion is conveyed by an Anglo-Saxon word, while the refined and more abstract meaning is expressed by a Norman-French term, which in turn is typically Latin-based, the difference between “pig” and “pork” being a case in point.

In the case of “freedom,” the Germanic word has a Franco-Latin counterpart in “liberty.” Generally speaking, “freedom” serves to convey the absence of physical constraint, as in talk about water flowing freely. Applied to human life, the absence of physical constraint may refer to objects not forming an obstacle for a human being’s actions, or to the absence of physical impediments brought to bear by other human beings on someone’s actions. By contrast, the presence or absence of the freedom called, in English, “liberty” turns not on natural but on normative, “moral” factors or considerations. In ancient parlance, “the moral” marks the opposite of the natural, as in “moral philosophy” as opposed to “natural philosophy,” or in “moral sciences,” meaning the humanities, as opposed to “natural science.” In this older and by now largely obsolete meaning the moral does not involve narrowly ethical concerns and regulations but comportmental controls and conditions of all kinds.

“Liberty” in the normative, “moral” sense then operates in a rule-governed context in which certain things may be permitted, others forbidden and yet others commanded. In such a context, “liberty” serves to single out those actions that are permitted, although not necessarily required. In particular, “liberty” refers to a sphere, range or kind of action to which an actor under a certain classification is entitled, either by exceptional privilege or by a more general prerogative. Grammatically considered, “liberty,” in its basically normative sense, is subject to pluralization, yielding “liberties” as the sum total of the specific forms of liberty possessed by a kind or a group of agents. By contrast, “freedom” is a *singulare tantum*. On the other hand, the noun “freedom” possesses a derivative adjective, “free,” which is not the case for “liberty.” Hence the typical recourse to “free” when an adjective correlated with “liberty” is needed.

The grammar and semantics of “liberty” first comes to the fore in the founding document of what was to develop, eventually, almost half a millennium later, into English liberalism (Whigs) and constitutionalism (Glorious Revolution, 1688), viz., the *Magna Carta Libertatum* or Great Charter of Liberties (1215). To be sure, this historical document is not the *Urtext* of modern, liberal democracy it sometimes is made out to be. The liberties in question pertained to those of the king’s subjects that were considered “free” (“freemen,” German *Freiherrn*) in the first place, essentially the barons of the realms. Still, subject to the subsequent extension of the free-men status to other segments of society, the Charter could provide a blueprint for later more broadly extended, even popular entitlements.

Among the liberties confirmed by a reluctant king to his free, baronial subjects, who had challenged him over withholding previously established entitle-

ments, were the protection from the undue search of house and property and the granting of a fair trial before peers. The notion of liberties employed in the document served to indicate not only a license to do or not do a certain thing on the part of the one being so entitled. It equally – even antecedently, logically speaking – served to indicate a matching binding force or obligation, not to say duty, on the part of the king (and his officers and agents) to abstain from interfering with, or infringing upon, his subjects' exercise of their liberties.

The intrinsic linkage between liberties and obligations is preserved in early modern thinking about the taming and framing of the emerging unlimited, “absolute” power of the sovereign, typically a dynastically entitled monarch but also admitting of a non-natural person (body of rulers) or even a popular sovereign, as in Rousseau's conception of the sovereignty of the people. The novel dimension that accrues to modern liberty is the focus on the conceptual space for and the actual extent of individual choice to be exercised under the liberties granted and guaranteed by the sovereign ruler. The modern individual is construed as being “at liberty” in the enjoyment of its entitlements, the latter increasingly understood as based on universal, “natural” norms (natural law, natural rights) that exceed, or rather precede, princely prerogative.

But the modern philosophical discourse about liberty adds not only a proto-liberal dimension to the traditional notions of liberty and liberties that run from Hobbes through Locke to J. St. Mill, with the latter drawing explicitly on W. v. Humboldt's earlier seminal essay on the issue (*Ideen zu einem Versuch, die Grenzen der Wirksamkeit des Staats zu bestimmen*; 1792). In addition, modern political philosophy retrieves and revives another, earlier line of thinking about liberty in a civic context. This political reading of liberty chiefly draws on the Roman republican conception of the state as a commonwealth established and maintained by public interest (*res publica*) rather than private pursuits and personal preferences. Roman-inspired republicanism had survived the end of the Republic and the eventual fall of the (Western) Roman Empire altogether in the political institution of independent, “free” city-size republics, to be found throughout medieval and early modern times in northern Italy and Imperial Germany.

In its strict variety, chiefly exemplified by Machiavelli, early modern neo-Roman republicanism seeks to emulate the Roman ethos of civic service (“virtue”), to the point of advocating anachronistic policies and practices. In an alternative development, early modern post-Roman republicanism arises as a historically informed but suitable adapted alternative to contemporary monarchical rule under the guise of anti-monarchical thinking and agitation (monarchomachism). Politically though, early modern republicanism risked to grow increasingly insignificant, theoretically as well as practically, due to the sweeping success of the system of absolute monarchy on the European Continent and became

obsolete with the constitutional curtailment of monarchical rule in England (and Scotland).

The philosophical fallout of the manifest divergence between ancient republicanism and early modern European political theory and practice is reflected by a conceptual separation effectuated already by Bodin and formalized by Kant. Intent on addressing the role and function of liberty in the modern polity, modern political thinkers come to distinguish between the outward, constitutional form of a state – which may be monarchical, aristocratic or democratic, according to the classical account (Plato, Aristotle) – and the actual mode of government operative in a given state, which may reflect, if not the outward form, then the functional mode or “spirit” (Kant) of a republic. The republican feature thereby introduced into the normative conception of the modern state as such, independent of its particular constitution, is the absence of arbitrary, tyrannical rule and the rule of law in its stead.

The separation of liberty in the political sense from its earlier exclusive linkage with republicanism makes liberty a generic, even defining feature of the modern state normatively considered. From Locke through Montesquieu and Rousseau to Kant, Fichte and Hegel and even beyond freedom *qua* liberty thus becomes the core concept of political philosophy, independent of the particular political persuasions and across a spectrum of political positions that reach from liberalism through nationalism to socialism and from monarchism through republicanism to democratism. In each case liberty, however construed in particular terms, is tied to laws, however specified, and hence distinguished throughout from license and arbitrariness.

The transformation of republicanism from a political philosophy favoring one constitutional kind over others to a principal position advocating political liberty is most visible in Montesquieu, whose work combines historical – in fact, world-historical – breadth and analytic perspicuity. In particular, Montesquieu ties the crypto-republican rule of law to the institutional separation of the legislative and executive functions of government, which he considers more likely to arise and more lasting to work in constitutional monarchies than in popular republics. For Montesquieu political liberty, rather than being tied to republics old and new, pertains to any system of rule in which laws, suitably chosen and equitably enacted, rule rather than the arbitrary will of a despotic regime.

The republican-inspired linkage of liberty and law effectuated by Montesquieu reaches definitional status when he outright identifies liberty in a civil-political regard with the citizens’ ability to act in accordance with the laws and not to be forced to act against them. On Montesquieu’s assessment, a lawful state – one in which the law rather than whim rules supreme – provides for its

citizens a condition of civic protection (“security”) that allows and enhances the flourishing of their private as well as public lives. To be sure, for Montesquieu, who is as much a crypto-republican as he is a proto-historicist, the laws to be obeyed to beneficial effect are not abstractly established rational rules but prudently practiced principles that reflect the cultural conditions, including their physical bases (“climate”), shaping the customs and codes of a given continent, country or city (“spirit of the laws”).

The more rigorous, rationalist version of Montesquieu’s linkage of law and liberty is to be found in Rousseau, whose neo-republican account of the civil-social contract bases political obligations on freely entered agreements among the citizens-to-be. But even Rousseau needs to acknowledge and admit the founding function of a legendary lawgiver, such as Lycurgus in Sparta, for assuring the trust that a people places in its customary laws. With regard to the linkage of law and liberty adumbrated by earlier neo-republicans and maintained by Montesquieu on general grounds, Rousseau construes the transition from the state of nature to the civil state as an equivalent exchange of wild, pre-civil freedom for suitably socialized, civil liberty. In his defense of the latter’s socio-cultural superiority over the former, Rousseau even goes so far as to allow, if not advocate, the use of compulsion for transforming a mere human being into a citizen (“forced to be free;” *Du Contrat Social*, Bk. 1., Chpt. 7).

3 Eleutheronomy

Kant participates in the progressive program of modern political philosophy to systematically link freedom and law. In general, he follows the proto- or crypto-republican strand of modern political thought of tying (political or civic) liberty to the rule of law. In particular, Kant distinguishes between a state’s “mode of government” (*Regierungsart*), which specifies the manner in which political power is organized and exerted, and a state’s “constitutional form” (*Staatsform*), which details the formal source of political authority (see AA 06: 340f.). On the republicanly minded political outlook maintained by Kant, a state may – indeed, should – operate in a republican manner, regardless of its outward constitution, which may be a monarchy (for linguistic reasons, Kant’s preferred term is “autocracy,” understood in a non-pejorative sense, see AA 06: 339), an aristocracy or even a democracy. For Kant, who here follows Montesquieu, the definitional feature of an inwardly republican state is the institutional and functional separation of the legislative power from the executive power.

Kant also follows the specifically modern republican view (though not shared by Montesquieu) that the ultimate source of political power is a (ficti-

tious) social contract entered into by originally free human beings who seek to form a civic community on the twin principles of liberty and equality (“lawful freedom,” “civil equality;” *gesetzliche Freiheit, bürgerliche Gleichheit*, see AA 06: 314). Yet while Rousseau maintains that the transition from the state of nature to the civil state involves an equivalent exchange of natural freedom for civic freedom, or liberty, Kant considers freedom to be realized only in the civic state. For Kant, the state of nature, while even including its own normative order (natural law), is not set up to assure everyone’s enjoyment of the naturally ordained rights. Kant also follows modern republican thought, especially Rousseau, in attributing to a people civically united (“commonwealth,” *gemeines Wesen*, see AA 06: 338) a set of intentions or a will of its own (“generally unified will,” *allgemein vereinigter Wille*; “unified popular will”; *vereinigter Wille des Volkes*, see AA 06: 313) geared toward the common interest, different both from the will of an individual and that of some subset of the collective community, both of which are in pursuit of particular interests.

But unlike ancient republicanism and its modern revival from Machiavelli to Rousseau, with its normative orientation toward the cultivation of a civic ethos of dedication and sacrifice (*virtù, vertu*), Kant’s jurally reduced republicanism transforms civic duties into legal obligations and makes freedom in its maximally generalized form (“general freedom,” *allgemeine Freiheit*, see AA 06: 323) the object and objective of political rule. More specifically, Kant defines an intrinsically republican state as that community in which the freedom, or rather liberty – the German “*Freiheit*” covering both concepts – of one is restricted by its compatibility with everyone else’s freedom, with the further provision that the general social compatibility is established “according to general laws.” For Kant the republican state, or the state reduced to its republican core, is the state of law and right (*Rechtsstaat*).

Kant’s definition of the (republican) state in terms of common civic liberty is adumbrated already in the *Critique of Pure Reason* (1781), where it serves to rationally reconstruct the idea of the “Platonic republic” (*Platonische Republik*, see A 316/B 372f.), and is fully formulated in the late *Metaphysics of Morals* (1797), where it furnishes the formal principle of (juridical) law under conditions of everyone’s equal freedom. In the latter work the firm focus on freedom *qua* liberty as the principle of political authority is paired with a matching focus on freedom as the principle of ethics and the basis of all ethical duties. To be sure, Kant’s parallel treatment of (juridical) law and ethics in the *Metaphysics of Morals*, which is already prepared in his earlier lectures on natural law from 1784 (*Naturrecht Feyerabend*), also includes the strict separation of law and ethics in terms of the kind and mode of lawgiving involved.

In particular, according to Kant's pioneering separation of law and politics from ethics, juridical duties concern only an agent's outward action (the external use of arbitrary choice or the elective will) and hence are limited to prescribing conformity to the law ("legality," *Legalität*, see AA 06: 214), while ethical duties command moral motivation and prescribe the agent's moral mind-set ("morality," *Moralität*, see AA 06: 214). Accordingly, for Kant juridical motivation is externally based, functioning essentially by means of punishment (and reward), while ethical motivation is internally based and operates by way of self-motivation.

The differences in mode of motivation notwithstanding, the parallels between the forms and functions of juridical and ethical lawgiving and law following in Kant are striking. To begin with, Kant joins law and ethics systematically and architectonically by treating the two as integral parts of a single "metaphysics of morals," which he takes to coincide, in extensional terms, with the non-empirical foundations ("metaphysical first principles") of "practical philosophy" *qua* "moral philosophy" (see AA 06: 214–221). In addition to pairing the two parts, he links them by way of a general, comprehensive introduction, in addition to providing individual introductions for each part. While the introduction common to the separately advanced foundations of law and ethics does not advance a material principle common to both domains of practical philosophy, it presents freedom as the primary feature of both law and ethics. In particular, both (juridical) law and ethics are shown to involve the giving of laws or the exercise of legislation (*Gesetzgebung*) with respect to freedom ("eleutheronomy," *Eleutheronomie*, see AA 06: 378).

Again, the laws thus introduced in law and ethics – in both cases not empirically based or naturally geared, but freedom-based and freedom-geared laws ("laws of freedom," see AA 06: 214) – are specifically different in terms of their manner of origination and mode of motivation. Yet the formal introduction of ethics under a general conception of the legislation of liberty ("eleutheronomy") constitutes an important advance in ethical theory and in metaethics. Since antiquity and up to Kant ethics had been construed in essentially consequentialist terms, with the end involved in ethical action typically termed "happiness" (*eudaimonia*), alternatively rendered as "the good life" (*das gelungene Leben*), to indicate the character of the ethical end as an achievement and accomplishment. Even Rousseau, whose political philosophy turns on freedom *qua* liberty as the supreme goal of civic life, bases his ethics on sentiment and sensibility (self-love, pity) rather than on a rationally formed rule of willing.

In rejecting consequentialism (along with sentimentalism) in ethics in favor of moral deontology, Kant not only removes any preconceived end from the principle of morality. He also limits the moral principle to establishing those laws and those laws only that reflect both the negative freedom of willing from par-

ticular, merely personal motives and, for that matter, from merely “subjective” principles of action and the positive freedom of following self-given laws only (“autonomy,” see AA 06: 383). As in the republican model underlying Kant’s radical revolution in ethics, freedom consists in following one’s own laws, rather than laws imposed by a foreign force, be the latter operative from within or from without (“heteronomy”). To be sure, in the case of ethics the self-given laws are not the peculiar laws of a particular ethical community, which would amount only to morals in the contingent sense of customs or *mores* (*Sitten*). For Kant ethical laws, insofar as they reflect the non-empirical, “metaphysical” principles of any and all morals (“metaphysics of morals”) are universal and hold for every human being considered as a moral agent. Strictly speaking, in a modal regard, ethical laws are even unlike principal juridical laws in that the former hold universally and necessarily (“a priori”) for all human beings *qua* practical rational beings, while the latter pertain to human beings *qua* citizens of a juridical commonwealth, no matter how limited (“state right,” *Staatsrecht*, see AA 06:311), extensive (“peoples’ right,” *Völkerrecht*; also “peoples’ state law,” *Völkerstaatsrecht*, see AA 06: 343 and AA 06: 311) or comprehensive (“cosmopolitan right,” *Weltbürgerrecht*, see AA 06: 352).

The juridico-political model underlying Kant’s critical ethics not only informs the linkage of law and liberty in the conception of moral autonomy and accounts for the marginal role of the freedom of choice and of elective will (*Willkür*) in an ethics oriented toward the functional equivalent of a republican ethos of obligation. The paralegal character of ethics in Kant also shows in the parallelism, established in the late *Religion Within the Bounds of Reason Alone* (1793), between the specifically “political commonwealth” (*politisches gemeines Wesen*, see AA 06: 95) that is the state normatively considered and the specifically “ethical commonwealth” (*ethisches gemeines Wesen*, see AA 06: 95) formed by ethical agents under moral laws, commonly termed “church” (*Kirche*). A further politico-ethical parallelism pertains to Kant’s moral depth psychology, with its distinction between an upper and lower faculty of desire modeled on the bicameral system of representation in modern constitutional monarchies and their division into a lower and upper chamber or house. In a further instance of functional transfer from the legal vehicle to the ethical tenor, this time drawing on the republican principle of the separation of political powers, Kant attributes to ethical volition, in addition to the legislative function involved in the (ethical) autonomy of the pure, purely moral will, the executive function of enacting the law by its own authority (“autocracy,” *Autokratie*, see AA 06: 383). In a similar vein, he pairs the specifically juridical instrument of outer constraint or compulsion (*Zwang*) with its ethical counterpart under the guise of “self-constraint” (*Selbstzwang*, see AA 06: 379f. note).

To be sure, none of the featured parallels, analogies or metaphors, individually or collectively considered, amount to a juridization of ethics in Kant. The distinction between the two kinds of the legislation of freedom remains valid and is further confirmed by the structural affinities involved. Still, the extent and depth of the affinities also attests to the formative function that (juridical) law, in its principal foundation, exercises for the grounding and shaping of an ethics of freedom in Kant. Drawing again on the letter, rather than the spirit, of Nietzsche's genealogical project, Kant's critical ethics can be characterized as a case of "internalization" by means of which the intersubjective relations from one sphere become compressed into internal, intrasubjective relations in the other sphere. But the juridical genealogy of the basic traits of Kant's ethics also suggests caution when treating Kant's ethics as an ethics of modern, individualized or "Protestant" subjectivity, as polemically portrayed by Hegel and emphatically embraced by others. In terms of its clandestine functional and structural origin in (liberal) law and (republican) politics, Kantian ethics is at least as much an ethics of universal, "catholic" concern for the egalitarian commonwealth of freely rational agents ("realm of ends," *Reich der Zwecke*).

Bibliography

- Zöllner, Günter (2010): "Autocracy. The Psycho-Politics of Self-Rule in Plato and Kant." In: Edmondo Balsamão Pires/Burkhard Nonnenmacher/Stefan Büttner-von Stülpnagel (Eds.): *Relations of the Self. Coimbra*: Coimbra University Press, pp. 385–404.
- Zöllner, Günter (2011): "Kant's Political Anthropology." In: *Kant Yearbook* 3, pp. 131–161.
- Zöllner, Günter (2015): *Res Publica. Plato's "Republic" in Classical German Philosophy*. Hongkong: The Chinese University Press. [North American distribution: Albany, NY: State University of New York Press, 2015.]
- Zöllner, Günter (2017): "Garantir le bien public et les droits privés" Emmanuel Kant et James Madison sur l'État républicain fédéral moderne." In: Sophie Grapotte/Mai Lequan/Lukas Sosen (Eds.): *Kant et les penseurs de langue anglophone. Mélanges en l'honneur de Jean Ferrari*. Paris: Vrin, pp. 275–292.
- Zöllner, Günter (2018): "Inborn Freedom.' Kant's Republicanism and Its Historico-Systematic Context." In: Violetta L. Waibel/Margit Ruffing/David Wagner (Eds.): *Nature and Freedom. Proceedings of the 12th International Kant Congress*. Berlin/Boston: De Gruyter, 1: pp. 693–709.

John Walsh

The Fact of Freedom: Reinhold's Theory of Free Will Reconsidered

Abstract: K.L. Reinhold advocates a theory of free will as the capacity to choose for or against the moral law. Reinhold's theory has often been accused of being psychologistic due to its alleged appeal to empirical facts of consciousness. This paper argues that instead of merely positing free will as a fact of consciousness, Reinhold provides an argument for free will as a necessary condition for moral responsibility. This sheds new light on the development of the concept of will in the wake of Kant's practical philosophy.

1 Introduction

In his *Letters on the Kantian Philosophy Volume II*, K.L. Reinhold presents a theory of free will in the spirit of Kant's own theory, though with an emphasis on choice, viz. our capacity to choose for or against the moral law. Scholarship on Reinhold has been growing and free will has been a popular topic among recent scholars. Reinhold's theory is often criticized for being psychologistic.¹ Faustino Fabbianelli claims that Reinhold had difficulty "extricating himself from a psychological conception of freedom because his discussion remains grounded in mere facts of consciousness" (Fabbianelli 2000, 441).² Georg Wallwitz accuses Reinhold of carrying out a "renaturalization of Kant's concept of freedom, which is located in the intelligible sphere" (Wallwitz 1999, 131).³ Marion

¹ Exceptions to the psychologistic interpretation of Reinhold's theory of free will are Martin Bondeli and Jörg Noller. Bondeli denies that Reinhold's theory "transforms moral freedom into a kind of psychological dependence" (Bondeli 2001, 248) and recognizes that Reinhold argues for freedom as a necessary condition for our awareness of the moral law as binding upon us (Bondeli 2018, 529–530). Noller acknowledges that Reinhold's theory concerns the "entire use of freedom" and that Reinhold's attempt to *realize* Kant's theory through first-order "volitional tendencies" structured in drives does not necessarily imply that Reinhold's theory merely concerns "moral psychology" (Noller 2015, 226).

² Fabbianelli goes on to assert that a critic could reject Reinhold's defense of free will on the basis of its "psychologism," and questions how it is possible to claim that man is free on the basis of one's own individual consciousness (Fabbianelli 2000, 442).

³ While Wallwitz does not explicitly charge Reinhold with psychologism, he maintains that Reinhold's claim that freedom is not an object of faith, but rather an object of knowledge "vio-

Heinz maintains that Reinhold “substitutes the conception of an anthropologizing moral psychology for Kant’s project of a metaphysics of morals” (Heinz 2012, 169). Similarly, Günter Zöllner claims “with Reinhold, practical philosophy remains in the domain of facts of moral consciousness and of their logical-analytical explication within the scope of *moral psychology*,” (Zöllner 2005, 75) and contends that Reinhold proceeds in the phenomenological style of the popular philosophers, connecting empirical results with their theoretical implications (Zöllner 2005, 87).⁴ However, not all commentators take a disparaging stance on Reinhold’s alleged psychologistic account of free will. Daniel Breazeale claims that for Reinhold, “the *evidence* for positing the same [absolute freedom; JW] is purely *empirical*, indeed, *psychological*, since the ‘facts’ in question are always facts of *consciousness*, allegedly available to everyone within inner experience” (Breazeale 2012, 112).⁵ Breazeale notes that Reinhold’s appeal to facts of consciousness can be seen as an anticipation of existentialist thinkers who take our phenomenological experience of free will to be fundamental to the human condition (Breazeale 2017). To be sure, on the one hand there is something off-putting about making philosophical claims on the basis of facts of consciousness, and a serious worry that this practice amounts to making bald assertions. On the other hand, phenomenological approaches to certain philosophical issues, or to philosophy in general, should not be rejected out of hand. However, rather than weigh in on the *merits* of Reinhold’s alleged grounding of our knowledge of free will in facts of consciousness, I contend that this is not the most charitable interpretation of Reinhold’s argumentative strategy after all, or, better yet, that the appeal to facts of consciousness tells only half the story.

This paper does not provide a full treatment of Reinhold’s use of facts of consciousness. Reinhold’s employment of facts of consciousness changed with his own philosophical developments and played a significant role in his theoretical project of the *Elementarphilosophie*.⁶ An account of the role of facts of con-

lates the principle of sufficient reason, according to which an object of empirical knowledge must always have knowable grounds” (Wallwitz 1999, 131).

⁴ Daniel Breazeale rightly notes that Zöllner’s position entails that Reinhold’s account “is not a contribution to philosophy at all, but rather to empirical psychology” (Breazeale 2012, 113).

⁵ For a defense of Reinhold’s supposedly psychologistic reliance on facts of consciousness see Breazeale 2012, 112–116.

⁶ For example, in the early incarnation of Reinhold’s *Elementarphilosophie* in *Contributions to the Correction of Previous Misunderstandings of Philosophers Volume I* (1790), the “fact of consciousness” was connected solely to the *Satz des Bewußtseyns*. It was not until his *On the Foundation of Philosophical Knowledge* (1791) that Reinhold ever used “facts of consciousness” in the plural (Lazzari 2004, 232n7). For a thoroughgoing treatment of Reinhold’s systematic developments between 1789 and 1803 see Bondeli 1995.

sciousness in Reinhold's theoretical philosophy would steer us too far afield from the specific task of reexamining the alleged psychologism of Reinhold's theory of free will.⁷ Furthermore, this paper neither defends Reinhold's claim that free will must be conceived as the capacity to choose for or against the moral law, nor does it treat Reinhold's dispute with Kant on the correct definition of free will. Instead, the scope of this paper is limited to clarifying Reinhold's argumentative strategy in his theory of free will and refuting some of the more egregious accusations regarding the supposed psychologism of Reinhold's account. I hope that this modest aim may shed new light on the development of the concept of will in Classical German Philosophy.

Part 1 examines Reinhold's account of free will and Reinhold's claim that we know that we are free as a fact of consciousness. In Part 2 I argue that this claim is not an empirical assertion, but follows from our consciousness of the moral law. In Part 3 I show that Reinhold argues by *reductio ad absurdum* that free will must include a capacity to freely act contrary to the dictates of the moral law. In Part 4 I consider two objections tied to Reinhold's supposed psychologism: Prauss' objection that Reinhold's account of the person undermines his account of free will and Zöller's objection that Reinhold's account is guilty of the homunculus fallacy.

2 Reinhold's Theory of Free Will and Facts of Consciousness

I will limit my treatment of Reinhold's theory of free will to the period between 1792 and 1797. Thus, I will be concerned with Reinhold's theory as it is presented in *Letters on the Kantian Philosophy Volume II* (1792), *Contributions to the Correction of Previous Misunderstandings of Philosophers Volume II* (1794), and "Some Remarks on the Concepts of Freedom of the Will posed by I. Kant in the Introduction to the Metaphysical Foundations to the Doctrine of Right" (1797). Although Reinhold addresses free will in his *Attempt at a New Theory of the Human Faculty of Representation* (1789), he had not yet divorced the will from practical reason and in that work considers only moral action to be absolutely free (*Attempt* pp. 571–572, *RGS* 1.362). It was not until 1792 that Reinhold conceived of the will as absolute self-activity independent of not only the demand

7 For an investigation of the method of Reinhold's early *Elementarphilosophie*, viz. the relation between reflection, intellectual intuition, and the fact of consciousness see Breazeale 2006.

of desire, but also of the demand of pure practical reason.⁸ From 1792–1797 Reinhold consistently conceives of free will as a capacity to determine oneself in accordance with either the demand of desire or the demand of reason, i.e. to determine oneself in accordance with or contrary to the moral law.

In his *Letters on the Kantian Philosophy Volume II*, Reinhold asserts that we have two fundamental drives: the selfish drive, which expresses the demand of desire, and the unselfish drive, which expresses the demand of the moral law (*LII* 181–184, *RGS* 2/2.134–136).⁹ The will is the capacity to determine oneself for or against the demand of the unselfish drive vis-à-vis. the satisfaction or non-satisfaction of the selfish drive (*LII* 183, *RGS* 2/2.135). That is, the will is the capacity to determine oneself for or against the moral law given the demand of desire. The two drives are connected in such a way that the demand of the unselfish drive, i.e. the demand of the moral law, always relates to the possible fulfillment of a given desire. Reinhold characterizes these two demands as occasioning grounds. Self-determination to the demand of the moral law or to the demand of desire constitutes a decision (*Entschluß*) whereby one of the occasioning grounds is made into a determining ground of volition. This decision is absolutely free.

Reinhold makes clear in “Letter Six” that his presentation of the fundamental concepts of ethics (including the concept of free will) is not based on metaphysics at all, but rather is based on the original and universal faculties of human cognition, which make themselves known through facts of consciousness (*LII* 180, *RGS* 2/2.133). Reinhold asserts as facts of consciousness: “*First*, that in volition as well as in involuntary desire, that drive which can be moved to action only by pleasure and displeasure is active; – *second*, that in volition, in addition to that drive, *reason* is also active in a particular way” (*LII* 244, *RGS* 2/2.173). By asserting these facts, Reinhold takes himself to have established the basic structure of agency in general, and volition in particular. The first fact establishes that in any action (i.e. in both volition and involuntary desire) the selfish drive is active. As has already been mentioned above, the selfish drive expresses the de-

⁸ For discussions of Reinhold’s separation of the will from reason see Lazzari 2004, 167–222 and Noller 2012. For a discussion of this aspect of Reinhold’s account of free will in relation to the imputability of immoral action and the dispute between Kant and Reinhold in 1797, see Noller 2019.

⁹ By 1794 Reinhold had abandoned the two fundamental drives, the selfish drive and the unselfish drive, characteristic of his theory in the second volume of the *Letters*. However, this revision did not touch his underlying position that the will is a capacity to choose to satisfy or not satisfy a demand of desire given the demand of the moral law. For relevant characterizations of the will from 1794 see Reinhold 2004, 143. For relevant characterizations from 1797 see Reinhold, *SWII* 372, *RGS* 5/2.143.

mand of desire, i.e. the demand grounded in pleasure and displeasure. Actions that are grounded only in desire, are not volitions, but are instinctual and animal (LII 246, RGS 2/2.174). The second fact establishes that in volition, reason is active by means of expressing the demand of the moral law through the unselfish drive. Reinhold goes on to assert that it is also a fact of consciousness that in volition the act of decision takes place as “the special act of our *I* (the person in us)” (LII 44–245, RGS 2/2.173). It is this latter fact that is the “genuine fact of freedom” (LII 192/279–80). According to Reinhold, this free self-determination “is completely comprehensible to me from its effects, through which it emerges among the facts of consciousness, and to this extent it is no object of *faith*, but rather an object of the most real *knowledge* for me” (LII 194/284). While it is undeniable that Reinhold does assert our knowledge of free will as a fact of consciousness, I argue that Reinhold takes himself to follow Kant’s claim in the 2nd *Critique* that our consciousness of free will follows directly from our consciousness of the moral law. Reinhold then offers an argument based on moral responsibility in order to establish that this freedom cannot consist solely in the capacity to obey the moral law, but must also include the capacity to choose against the moral law.

3 Reexamination of Freedom as a Fact of Consciousness

A number of commentators interpret Reinhold’s claim that we are aware of our freedom as a fact of consciousness to be both immediate and empirical. Breazeale attributes to Reinhold the view that our free will “is something that we *immediately* and *directly* know about ourselves” (Breazeale 2012, 94). Pierluigi Valenza claims that freedom “is a fact that emerges directly in consciousness such that the experience of freedom is an object of psychology and anthropology” (Valenza 2012, 359). Similarly, Alessandro Lazzari claims that Reinhold “by no means held that the reality of freedom is to be derived from our consciousness of the moral law” (Lazzari 2004, 305n33). Yet Reinhold explicitly states that we are aware of our freedom by virtue of our awareness of the moral law: “the claim ‘that the concept of freedom first receives its reality through consciousness of the moral law’ from the *Critique of Practical Reason* is incontestably true” (LII 276, RGS 2/2.190).¹⁰ In “Some Remarks” Reinhold maintains that his

¹⁰ In *Contributions II* Reinhold claims that we know that the independence of the act of decision from the demand of desire is not an illusion “from the *consciousness* of the unique *law* that

concept of freedom is “drawn solely from consciousness of the moral law itself, from the *categorical imperative* alone” (SWII 393, RGS 5/2.150). Most scholars have overlooked these telling passages. Interestingly, Lazzari addresses the passage from Reinhold’s *Letters II*, but denies that Reinhold affirms Kant’s claim that our consciousness of freedom follows from consciousness of the moral law. As far as I know, Lazzari is the only recent scholar to take note of either of these passages and so I will focus on his position.

The issue is complicated by the fact that, in addition to our consciousness of the moral law, Reinhold seems to introduce another condition necessary for our consciousness of freedom. It will be necessary to cite the relevant passage from *Letters II* in full in order to clarify Reinhold’s position:

The claim ‘that the concept of freedom first receives its reality through consciousness of the moral law’ from the *Critique of Practical Reason* is incontestably true. The person can become conscious of the capacity to determine himself only insofar as he is conscious of the capacity to determine himself according to two different laws, and consequently insofar as he is conscious of these different laws themselves. But precisely for that reason freedom can also by no means consist in the capacity to follow only one of the two laws, and that *Kantian* claim can by no means have the meaning: ‘that the reality of freedom depends upon consciousness of the moral law *alone*.’ (LII 276, RGS 2/2.190)

In addressing Lazzari’s position, I will first refute his claim that Reinhold is criticizing Kant in the passage above and will then suggest that C.C.E. Schmid is a more likely candidate for the target of Reinhold’s remarks. Lazzari correctly notes that Reinhold is referring to Kant’s doctrine of the moral law as the *ratio cognoscendi* of freedom. Lazzari argues:

This assertion by Kant and the doctrine, which is immediately connected to it, of the moral law as the *ratio cognoscendi* of freedom is not to be understood [according to Reinhold; JW] such that consciousness of the moral law is the sole sufficient condition of consciousness of our own freedom, but rather, that it merely concerns a necessary condition. For in addition to consciousness of the moral law, that consciousness of the natural laws of desire must also be postulated as a necessary condition. What here seems to be a recommendation for a reading of the passage from the *CPrR* is in reality a criticism of Kant. For not only does Kant never mention a further condition that must be fulfilled in addition to consciousness of the moral law in order to accept the reality of free will, on the contrary his remarks support – as in the case of the famous footnote at the beginning of the Preface in the *CPrR* – the reading whereby consciousness of the moral law is the sole sufficient condition of our consciousness of freedom. (Lazzari 2004, 308–309)

in acts of the will announces itself before decision, is thought through the *ought*, and is called the *moral law* or the *law of the will*” (Reinhold 2004, 138).

Lazzari's point seems to be that Reinhold's supposed introduction of an additional condition necessary for consciousness of freedom is inconsistent with Kant's claim that consciousness of the moral law is the "sole sufficient condition," i.e. the necessary and sufficient condition, of consciousness of freedom. Lazzari would be correct if Reinhold did in fact introduce an additional necessary condition. However, Lazzari fails to recognize that Reinhold considers consciousness of the demand of desire to be entailed by consciousness of the moral law.

According to Reinhold, consciousness of the demand of desire is a necessary but not sufficient condition for a state of reflection, or circumspection (*Besonnenheit*), in which one can become conscious of the demand of the moral law (LII 305–306, RGS 2/2.205). Thus, consciousness of the moral law entails consciousness of the demand of desire. With respect to the passage in question from the *Letters II*, Reinhold is not really positing an additional condition because consciousness of the demand of desire is already entailed by consciousness of the moral law. That is not to say, of course, that the selfish drive is somehow contained in the unselfish drive. Reinhold considers these two drives to be original and independent. The point is rather that consciousness of the demand of the former is entailed by consciousness of the demand of the latter. Given that this is the case, Reinhold can still maintain the Kantian claim that consciousness of the moral law is necessary and sufficient for consciousness of freedom. That Reinhold is not, as Lazzari supposes, considering consciousness of the moral law to be only a necessary condition is also supported by Reinhold's reference of a passage where Kant asserts the sufficiency of consciousness of the moral law for consciousness of freedom: "It is therefore the *moral law*, of which we become immediately conscious (as soon as we draw up maxims of the will for ourselves), that *first* offers itself to us and [...] leads directly to the concept of freedom" (CPrR 5:29–30).¹¹ This is surely the passage Reinhold has in mind when he asserts that "the claim 'that the concept of freedom first receives its reality through consciousness of the moral law' from the *Critique of Practical Reason* is incontestably true" (LII 276, RGS 2/2.190). Instead of criticizing Kant, it is more likely that Reinhold is criticizing Schmid's interpretation of Kant.

Just a few pages before the passage we have been considering from *Letters II*, Reinhold refers to C.C.E. Schmid's *Lexicon for the Easier Use of the Kantian Writings*:

¹¹ Lazzari overlooks this passage and restricts his discussion of Kant to the Preface of the *Critique of Practical Reason*.

On *practical freedom*, and thus certainly on freedom of the will, the same *Lexicon* reads: ‘In the *positive* sense it is dependence of the will upon reason which determines it immediately, upon the pure moral law, the autonomy of the will;’ whereby p. 59 and 238 of the *Critique of Practical Reason* are cited. (*LII* p. 271, *RGS* 2/2.187)¹²

Schmid’s explicit characterization of freedom as “dependence” (*Abhängigkeit*) on reason and the moral law, as well as Schmid’s reference to the *Critique of Practical Reason* make him the likely target when Reinhold asserts “that *Kantian* claim can by no means have the meaning: ‘that the reality of freedom depends upon consciousness of the moral law *alone*’” (*LII* 276, *RGS* 2/2.190). As we will see in the next section, Schmid locates freedom wholly in the intelligible realm, thereby precluding a role for rational deliberation about our sensible desires in his account of free agency. Accordingly, Reinhold is clarifying that the moral law is something we experience as the normative constraint of our sensible desires and it applies to the adoption of particular maxims, which always have an empirical content (even if, as Kant maintains, they are not adopted *by virtue of* that content, but rather, in the case of moral actions, by virtue of their form). Before turning to Schmid and Reinhold’s argument for free will as embedded in his response to Schmid, I would like to make a few additional remarks on Reinhold’s alleged grounding of free will in empirical facts of consciousness.

To be fair to those interpreters who take the fact of freedom to be immediate and empirical, as we saw in Section 1.1, Reinhold does claim that the original faculties of the mind (one of which he considers to be free will) make themselves known through facts of consciousness and that he has based the results of his investigation on facts of consciousness. Indeed, Reinhold admits that his claims are “unargued” and “unproven.” However, I contend that Reinhold’s position on the foundation of his own claims is the result of his lack of a systematic integration for the same. Reinhold promises a “future, more precise argument” for his basic concepts of ethics and natural right (*LII* 181, *RGS* 2/2.134), and mentions a “future system of pure ethics and natural right” (*LII* 179, *RGS* 2/2.132). Of course, Reinhold also envisioned that this system of ethics and natural right would be part of a complete *Elementarphilosophie* that united theoretical and practical philosophy in a comprehensive system:

Only a *single* system can be built upon fully determinate fundamental concepts and only a *single philosophy* is possible that in its principles is the correct expression of the original

12 Cf. Schmid 1788, 179.

arrangement of our faculty of cognition and our faculty of desire, or of the necessary and universal law, to which the human spirit by its nature is bound. (*LII* 21, *RGS* 2/2.21)

Thus, Reinhold's admission that his concepts are unargued and unproven can be seen to stem from his lack of arguments that would incorporate these concepts into an all-encompassing philosophical system. It is not the case that Reinhold has no arguments whatsoever for his concepts. As we will see, Reinhold argues for his concept of free will as a necessary condition for moral responsibility.

4 Reinhold's Argument

Reinhold argues for the concept of free will, viz. the capacity for spontaneous self-determination for or against the moral law, from the premise that free will is a necessary condition for moral responsibility. His argument is imbedded in his response to the conception of free will espoused by C.C.E. Schmid. I will briefly discuss Schmid's conception in order to bring Reinhold's argument to light.

In his *Attempt at a Moral Philosophy*, Schmid asserts, in what he takes to be the spirit of Kant's own theory of freedom, the doctrine of intelligible fatalism, i.e. that all actions are determined by intelligible causality. Actions with immediate determining grounds in sensible, temporal circumstances (*sinnliche Zeitumstände*) are ultimately grounded in an *intelligible Naturfatalismus*, "i.e. the assertion of the natural necessity of all actions of a rational being according to laws of the causality of things in themselves" (Schmid 1790, 211). With respect to morality, Schmid thought it was enough to believe that *Zeitumstände* cannot compel us to act irrationally (i.e. immorally) and that pure reason can provide a determining ground for actions through the moral law (Schmid 1790, 198 ff). According to Schmid, both immoral action and non-moral action (action with no moral significance) are grounded in an intelligible causality underlying objects of experience, and moral action is grounded in an intelligible causality underlying the agent as a thing in itself, i.e. moral action is grounded in the intelligible causality of pure practical reason. Schmid restricts freedom to moral action and claims that immoral action is the result of a hindrance to the efficacy of reason.

Reinhold notes that identifying free will solely with reason's self-legislation of the moral law entails the impossibility of free immoral action (*LII* 267, *RGS* 2/2.185). Because Schmid explicitly admits that freedom of the will is restricted to moral action, merely stating that a conception of free will restricted to moral action precludes freedom for immoral action is not sufficient to refute Schmid's conception of free will. Reinhold needs to show that this conception is inherently

problematic. Reinhold concedes that at least Schmid is consistent enough to recognize that if free will is restricted to moral action, then the ground of immoral action must “be sought outside of the will in external obstacles” (*LII* 296, *RGS* 2/2.200). However, if the ground of immoral action is posited in external obstacles, then the ground of moral action would consist in the absence of such obstacles:

Moral actions would inevitably take place through a completely involuntary activity of practical reason *as long as there was no obstacle there*; and both moral and immoral actions would thus have to be attributed solely to the presence or absence of such an obstacle. (*LII* 296–297, *RGS* 2/2.200).¹³

Of course, Reinhold’s argument as it stands is rather underdeveloped. It only establishes that *if* we are morally responsible for our actions, then a conception of free will that is restricted to moral actions would undermine that moral responsibility insofar as immoral actions would be directly grounded in the presence of obstacles hindering the efficacy of pure practical reason and moral actions would ultimately be grounded in the absence of such obstacles. The argument is a nonstarter if the antecedent is not established, i.e. it must be established that we really are morally responsible for our actions.

Reinhold maintains that “the moral law is absolutely necessary” (*LII* 196, *RGS* 2/2.144) and that the moral law is “simply given to it [the will; JW] by *pure reason*” (*LII* 285–286, *RGS* 2/2.194–195). However, these are not bald assertions. Reinhold takes himself to be following Kant: “*Kant* was the first to demonstrate that *pure reason* is self-active in moral lawgiving and that the law set forth solely by it is the objective determining ground of a moral action” (*LII* 304, *RGS* 2/2.204). Admittedly, Reinhold seems to presuppose that consciousness of the moral law as binding entails consciousness that we are morally responsible for our actions. This assumption is most clear in “Some Remarks,” where Reinhold asserts an analytic connection between the morality of an action and its imputability (*Zurechnungsfähigkeit*) (*SWII* 364, *RGS* 5/2.141).¹⁴ With this assumption, we can reconstruct Reinhold’s argument. Given that we are bound by the moral law, we are morally responsible for our actions. Suppose that we were only free to obey the moral law. Then moral responsibility would be undermined, because, as we saw in Reinhold’s treatment of Schmid, immoral actions would be grounded in obstacles preventing moral action, and moral action would in turn be grounded in the absence of those obstacles. If moral responsibility were under-

¹³ For Reinhold’s employment of this argument in *Contributions II* see Reinhold 2004, 136, 141. Prauss recognizes that this is crucial to Reinhold’s argumentative strategy (Prauss 1983, 86).

¹⁴ I am not sure Reinhold is wrong to think this.

mined, then so too would our being bound by the moral law. Our consciousness of the moral law is a priori, and therefore necessary. Accordingly, the supposition that we are only free to obey the moral law contradicts our a priori consciousness of the moral law as binding upon us. The contradiction is lifted if we grant our freedom to transgress the moral law. In this way, Reinhold's concept of freedom as the capacity to choose for or against the moral law is established by accepting the result of the *Critique of Practical Reason* that we are a priori conscious of the moral law as binding upon us, and then, given the premise that moral obligation entails moral responsibility, arguing by *reductio ad absurdum* that we must also be free to transgress the moral law.

5 Refutation of the Objections Concerning Reinhold's Conception of the Person

Reinhold's conception of the person has come under fire recently from several commentators, most notably Gerold Prauss and Günter Zöller. Given the significance of this conception for Reinhold's account of free will, it is incumbent upon us to consider these objections. Prauss questions the intelligibility of Reinhold's account of self-determination given that Reinhold's employment of "person" is inconsistent. Prauss' objection can be dealt with rather swiftly. It is not my intention to create a straw man. Given the stature of Prauss as a Kant scholar and his connection to scholarship on Kant and Reinhold on free will, it is important to consider Prauss' objection despite its shallow import.¹⁵ Zöller raises concerns that Reinhold's account of the person as the locus of decision in intentional action is guilty of the homunculus fallacy. Although Reinhold's account need not necessarily lead to an infinite regress, there are concerns connected to his postulation of a "person in us" that performs the self-determining act of decision.

Gerold Prauss argues that Reinhold's conception of free will is untenable because his account of the unity of the subject, or the person, falls asunder (Prauss 1983, 90). According to Prauss, Reinhold sometimes identifies the "person qua person" (*Person als Person*) with the necessary self-activity of practical reason, and sometimes identifies the person with the will and its free self-determination. This is supposedly problematic insofar as it is precisely the *distinction* between the will and practical reason that is so fundamental to Reinhold's account. The contradiction between these notions of "person" and the consequent breakdown

¹⁵ For a discussion of Prauss with respect to Kant and Reinhold on free will see: Allison 1990, 134–135; Ameriks 2012, 76; Noller 2015, 37, 208; and Zöller 2005, 86.

of the distinction between will and practical reason, which is essential to Reinhold's theory of free will, allegedly amounts to the breakdown of Reinhold's entire account of the will. Prauss is correct that Reinhold sometimes conflates two conceptions of "person." However, he is incorrect in properly identifying the conflation. While Reinhold discusses the person with respect to the self-activity of practical reason in some passages and discusses the person with respect to the will and its freedom of self-determination in others, Reinhold's conceptions of the person seem to be: (1) a generic notion for the individual and all his constituent faculties; and (2) the subject of mental states. Although Reinhold does often conflate these conceptions, I contend that this is not ultimately problematic and that Prauss has thrown the baby out with the bathwater.

Reinhold's lack of consistency in his use of the term "person" does not necessarily undermine the claim that in any act of volition, there must be a constituent intentional act of decision whereby the subject determines itself to a particular course of action. There are numerous passages where Reinhold employs "person" as the self, i.e. the subject of decision (*Entschluß*) constituting an act of volition.¹⁶ There are indeed other passages where he seems to employ "person" not merely as the subject (of decision or of particular mental states), but rather as a general term for the individual.¹⁷ In the latter usage, "person" is an aggregate term for all of the individual's faculties, including those whose activities are not intentional acts of the self-conscious subject, but rather sub-conscious activities, the effects of which are presented *to* the self.¹⁸ The inconsistency in these two uses is not necessarily indicative of a doctrinal contradiction. It simply does not follow from terminological carelessness that Reinhold's entire account of free will falls apart. There is no contradiction in maintaining a conception of an individual as the aggregate of his constituent faculties while also maintaining a conception of a subject of mental states. Indeed, there is no substantive difficulty in quickly resolving this apparent semantic oversight. Giving one or the other a different name would alleviate this superficial incon-

16 See Reinhold, *LII* 184, *RGS* 2/2.136; *LII* 185, *RGS* 2/2.137; *LII* 187, *RGS* 2/2.139; *LII* 188, *RGS* 2/2.139; *LII* 189, *RGS* 2/2.140; *LII* 197, *RGS* 2/2.144; *LII* 207, *RGS* 2/2.151; *LII* 210, *RGS* 2/2.153; *LII* 211, *RGS* 2/2.154; *LII* 215, *RGS* 2/2.156; *LII* 215, *RGS* 2/2.157.

17 See Reinhold, *LII* 69, *RGS* 2/2.54; *LII* 182, *RGS* 2/2.134; *LII* 183, *RGS* 2/2.135; *LII* 185, *RGS* 2/2.137; *LII* 186, *RGS* 2/2.138; *LII* 189, *RGS* 2/2.140; *LII* 191, *RGS* 2/2.141; *LII* 201, *RGS* 2/2.148; *LII* 203, *RGS* 2/2.148; *LII* 203, *RGS* 2/2.149; *LII* 207, *RGS* 2/2.151.

18 Reinhold considers the activities of the faculty of desire and practical reason to be *for* the person qua intentional subject: "The *demands* of desire and the *demand* of practical reason are interrelated insofar as in volition both are directed at the person as the subject of freedom" (Reinhold 2004, 168). He also considers these activities to be involuntary (*LII*, 182, *RGS*, 2/2.134).

sistency. In this way, it is clear that in drawing attention to Reinhold's conflation of terms, Prauss has conflated a terminological inconsistency with a doctrinal contradiction.

Günter Zöller claims that Reinhold's account of the will is guilty of smuggling a homunculus into his theory of subjectivity (Zöller 2005, 82).¹⁹ The suspicion is that Reinhold's installation of a faculty of choice (*Willkür*) in the individual, which is responsible for choosing to act for or against the moral law, is tantamount to the homunculus fallacy, viz. an infinite regress of "little men" postulated as a locus of decision in an individual. Daniel Breazeale argues that Zöller's concern is really based on the "*incomprehensibility* of a radically free choice on Reinhold's account" (Breazeale 2012, 108–109). The suspicion that Reinhold's account of free will amounts to an infinite regress of free acts of decision made by homunculi within us may be the vestige of the reluctance to accept at face value Reinhold's assertion that the ground of a free act is freedom itself (*LII* 282, *RGS* 2/2.193), i.e. that freedom is simply its own self-contained ground. Given that Reinhold postulates the free act as a first cause, Breazeale is correct that there is no need to posit an infinite regress of acts. Nevertheless, some of Reinhold's characterizations of the "person" resemble the postulation of a homunculus. While Zöller does not refer to any specific passages from Reinhold, consideration of Reinhold's problematic language is certainly worthwhile given the centrality of the person in his account of free agency.

As I already had occasion to cite, Reinhold refers to the act of decision as "the special act of our *I* (the person in us)" (*LII* 173/245). This language of a "person in us" might well raise concerns that Reinhold's account of the person resembles a homunculus. I admit that Reinhold's phrasing may be guilty of suggesting this. I propose a charitable interpretation of what Reinhold might mean by "the special act of our *I*." The problem is greatly mitigated if by "the *I*" and "the person in us," we take Reinhold not to assign metaphysical significance, but rather to assert a necessary condition for the intentional act of self-determination, namely that it include consciousness of the demands of desire and of the moral law. On this reading, the "person in us" is no metaphysical postulation; rather, it merely designates consciousness as a necessary condition for self-determination. In this way, Reinhold's language can, at least in part, be seen as a response to C.C.E. Schmid's conception of free will. Whereas Schmid posited free agency as a noumenal activity outside the purview of possible phenomenal consciousness, Reinhold considered it necessary for the free act of decision that the subject be conscious of the demands of desire and of the moral law. Of

19 For a reiteration of this charge see Noller 2015, 234.

course, the I is not conscious of the actual performance of the decision, but is conscious of the effect. The free activity itself is indeed intelligible; however, our phenomenal consciousness is a necessary condition for the exercise of this capacity. The intelligible activity of freedom is incomprehensible for beings such as ourselves; nevertheless, we must presuppose that it is an activity of the I, i.e. of the subject of transcendental faculties: “The *subject* of the transcendental faculties is likewise the subject of the empirical faculties if those faculties are not to be *transcendent*, but rather *transcendental* – i.e. related *a priori* to the empirical” (SWII 393, RGS 5/2.151). By bridging the gap excavated by Schmid’s intelligible fatalism, Reinhold believed himself to have delivered the determinate concept of free will, which Kant could only prepare. Whether Reinhold succeeded in this ambitious enterprise is a question that must be answered on another occasion.

6 Conclusion

Although Reinhold certainly thought that the reality of freedom was something that could be recognized as a fact by even the most common understanding, he is not guilty, as he has often been accused, of grounding knowledge of free will in mere empirical facts of consciousness. Instead, Reinhold assumes the givenness of the moral law as binding upon us and argues by *reductio ad absurdum* that we must be free to choose to act in accordance with this law or to transgress it. Reinhold’s theory of free will is inextricably connected to Kant’s theory of the same. While the relation between these two theories has previously been explored, I have argued that Reinhold appeals to Kant in his attempted demonstration of free will more than is generally acknowledged. It is my hope that by clarifying Reinhold’s argumentative strategy, we might better situate his theory in the context of the development of the concept of will in Classical German philosophy.

Bibliography and Abbreviations

All translations of Kant’s works refer to the *Cambridge Edition of the Works of Immanuel Kant*. Passages from Kant’s works are cited according to the volume and pagination of *Kants Werke* (Akademie-Textausgabe). Unaltered photocopy reprint of the text from the publication series of Kant’s complete works initiated by the Prussian Academy of Sciences 1900 ff. 29 vols., Berlin: De Gruyter). Translations of Schmid (1790) are co-translated by Jörg Noller and me, and are taken from the volume under contract with Cambridge University Press, *Kant’s Early Critics on Freedom of the Will*. All other translations are my own.

Attempt	Attempt at a New Theory of the Human Faculty of Representation
CPrR	Critique of Practical Reason
LII	Letters on the Kantian Philosophy Volume II
RGS	Karl Leonhard Reinhold Gesammelte Schriften
SWII	Selected Writings Volume II

- Allison, Henry (1990): *Kant's Theory of Freedom*. Cambridge: Cambridge University Press.
- Ameriks, Karl (2012): "Ambiguities in the Will: Reinhold and Kant, *Briefe II*." In: Bondeli, Martin/Heinz, Marion/Stolz, Violetta (Eds.): *Wille, Willkür, Freiheit: Reinholds Freiheitskonzeption im Kontext der Philosophie des 18. Jahrhunderts*. Berlin: De Gruyter, pp. 71–89.
- Bondeli, Martin (1995): *Das Anfangsproblem bei Karl Leonhard Reinhold: Eine systematische und entwicklungsgeschichtliche Untersuchung zur Philosophie Reinholds in der Zeit von 1789–1803*. Frankfurt am Main: Vittorio Klostermann.
- Bondeli, Martin (2001): "Freiheit im Anschluss an Kant. Zur Kant-Reinhold Kontroverse und ihren Folgen." In: Horstmann, Rolf-Peter/Schumacher, Ralph/Volker, Gerhard (Eds.): *Kant und die Berliner Aufklärung: Akten des IX. Internationalen Kant-Kongresses*. Berlin: De Gruyter, pp. 243–251.
- Bondeli, Martin (2018): "Freiheit, Gewissen und Gesetz. Zu Kants und Reinholds Disput über Willensfreiheit." In: Ruffing, Margit/Waibel, Violetta L. (Eds.): *Natur und Freiheit: Akten des XII. Internationalen Kant-Kongresses*. Berlin: De Gruyter, pp. 517–532.
- Breazeale, Daniel (2006): "Reflection, Intellectual Intuition, and the Fact of Consciousness: Remarks on the Method of Reinhold's *Elementarphilosophie* (1789–1791)." In: Valenza, Pierluigi (Ed.): *K. L. Reinhold am Vorhof des Idealismus*. Pisa, Rom: Istituti editoriali e poligrafici internazionali, pp. 35–57.
- Breazeale, Daniel (2012): "The Fate of Kantian Freedom: One Cheer (More) for Reinhold." In: Bondeli, Martin/Heinz, Marion/Stolz, Violetta (Eds.): *Wille, Willkür, Freiheit: Reinholds Freiheitskonzeption im Kontext der Philosophie des 18. Jahrhunderts*. Berlin: De Gruyter, pp. 91–123.
- Breazeale, Daniel (20 April 2017): "The Primacy of the Practical and the Freedom of the Will." Invited Lecture, Leuven Research Group in Classical German Philosophy, Leuven.
- Fabbianelli, Faustino (2000): "Die Theorie der Willensfreiheit in den 'Briefe über die Kantische Philosophie' von Karl Leonhard Reinhold." In: *Philosophisches Jahrbuch* 107. No. 2, pp. 428–443.
- Heinz, Marion (2012): "Moralpsychologie statt Metaphysik der Sitten." In: Bondeli, Martin/Heinz, Marion/Stolz, Violetta (Eds.): *Wille, Willkür, Freiheit: Reinholds Freiheitskonzeption im Kontext der Philosophie des 18. Jahrhunderts*. Berlin: De Gruyter, pp. 167–191.
- Kant, Immanuel (1996): *CPrR* | "Critique of Practical Reason (1st ed. 1788)." In: *Practical Philosophy*. Gregor, Mary J. (trans.). *The Cambridge Edition of the Works of Immanuel Kant*. Paul Guyer/Allen Wood (Eds.). New York, Cambridge: Cambridge University Press, pp. 133–271.
- Lazzari, Alessandro (2004): *Das Eine, was der Menschheit Noth ist. Einheit und Freiheit in der Philosophie Karl Leonhard Reinholds (1789–1792)*. Stuttgart-Bad Cannstatt: Frommann Holzboog.

- Noller, Jörg (2012): "Die Praktische Vernunft ist kein Wille. Reinholds personalitätstheoretische Kritik der Kantischen Freiheitslehre." In: Bondeli, Martin/Heinz, Marion/Stolz, Violetta (Eds.): *Wille, Willkür, Freiheit: Reinholds Freiheitskonzeption im Kontext der Philosophie des 18. Jahrhunderts*. Berlin: De Gruyter, pp. 193–221.
- Noller, Jörg (2015): *Die Bestimmung des Willens. Zum Problem individueller Freiheit im Ausgang von Kant*. Freiburg, München: Verlag Karl Alber.
- Noller, Jörg (2019): "'Practical Reason is not the Will': Kant and Reinhold's Dilemma." In: *European Journal of Philosophy*. <https://doi.org/10.1111/ejop.12448>.
- Prauss, Gerold (1983): *Kant über Freiheit als Autonomie*. Frankfurt am Main: Vittorio Klostermann.
- Reinhold, Karl Leonhard (2004): *CII* | "Beiträge zur Berichtigung bisheriger Mißverständnisse der Philosophen. Zweiter Band [Contributions to the Correction of Previous Misunderstandings of Philosophers Volume II] (1st ed. 1794)." Fabbianelli, Faustino (Ed.). Hamburg: Felix Meiner.
- Reinhold, Karl Leonhard (2008): *LII* | "Briefe über die Kantische Philosophie Zweyter Band [Letters on the Kantian Philosophy Volume II] (1st ed. 1792)." In: *Karl Leonhard Reinhold Gesammelte Schriften*. Vol. 2/2. Martin Bondeli (Ed.). Basel: Schwabe Verlag.
- Reinhold, Karl Leonhard (2013): *Attempt* | "Versuch einer neuen Theorie des menschlichen Vorstellungsvermögens [Attempt at a New Theory of the Human Faculty of Representation] (1st ed. 1789)." In: *Karl Leonhard Reinhold Gesammelte Schriften*. Vol. 1. Martin Bondeli (Ed.). Basel: Schwabe Verlag.
- Reinhold, Karl Leonhard (2017): *SWII* | "Auswahl vermischter Schriften Zweyter Teil [Selected Writings Volume II] (1st ed. 1797)." In: *Karl Leonhard Reinhold Gesammelte Schriften*. Vol. 5/2. Martin Bondeli/Silvan Imhof (Eds.). Basel: Schwabe Verlag.
- Schmid, Carl Christian Eerhard (1788): *Wörterbuch zum leichtern Gebrauch der Kantischen Schriften [Lexicon for the Easier Use of the Kantian Writings]*. Jena: Cröker.
- Schmid, Carl Christian Eerhard (1790): *Versuch einer Moralphilosophie [Attempt at a Moral Philosophy]*. Jena: Cröker.
- Valenza, Pierluigi (2012): "Reinholds Freiheitskonzeption im Atheismusstreit." In: Bondeli, Martin/Heinz, Marion/Stolz, Violetta (Eds.): *Wille, Willkür, Freiheit: Reinholds Freiheitskonzeption im Kontext der Philosophie des 18. Jahrhunderts*. Berlin: De Gruyter, pp. 349–369.
- Wallwitz, Georg (1999): "Fichte und das Problem des intelligiblen Fatalismus." In: Hammacher, Klaus/Schottky, Richard/Schrader, Wolfgang H. (Eds.): *Transzendente Logik, Fichte Studien* 15, pp. 121–145.
- Zöllner, Günter (2005): "Von Reinhold zu Kant. Zur Grundlegung der Moralphilosophie zwischen Vernunft und Willkür." In: *Archivio di Filosofia (K. L. Reinhold. Alle soglie dell'idealismo)* 73, nos. 1–3, pp. 73–91.

Amit Kravitz

On the Real Possibility of a Pure Moral Will: Maimon vs. Kant

Abstract: One of the main challenges of Kant's moral theory is to show why a finite agent, whose will is already affected by natural driving-forces which generate pleasure, would actually choose to subordinate them to a pure driving-force – 'respect for the moral law' – which generates no pleasure (and even humiliates, as Kant has it). After outlining the subtle way Kant tackles this challenge I focus on Salomon Maimon's alternative. Contrary to post-Kantian philosophers like Fichte or Schelling, Maimon does not try to trace back the 'fact of reason' – morality's point of departure – to a prior metaphysical background. Rather, this 'fact' can be seen according to Maimon as an outcome of an observable natural driving-force (however a unique one). I then present at length Maimon's alternative to Kant concerning the structure of a finite will and the possibility of a moral action and assess the strengths and weaknesses of Maimon's solution, as well as its implication on his moral theory as a whole.

Es liegt also der moralische Werth der Handlung nicht in der Wirkung, die daraus erwartet wird (Kant, GMS, AA 04: 401.3f.).

Die Tugend gewährt die reinste immer dauernde und beständig zunehmende Glückseligkeit (Maimon, Ueber die ersten Gründe der Moral, 476).

1 The Problem

As Kant stresses time and again,¹ 'ought to' implies 'can.' Evidently, what is at stake is the question concerning the possibility of morality, i.e. the possibility of a finite will to be determined by a pure moral driving-force.

However, the term 'possibility' in the Kantian moral context cannot be understood merely logically; in Kant's morality it obtains an inherent connection to *finitude*. 'Imperative,' for instance, is brought to the fore only with reference to a finite will – that is, with reference to a will which is affected (however not yet fully determined²) by definition also by driving-forces (*Triebfeder*) other than the respect for the moral law (put differently, for a divine will, which is de-

¹ See, for instance, KpV, AA V: 30; RGV, AA VI: 45; TP, AA VIII: 284.

² Cf. GMS, AA IV: 387.

terminated “by itself [von selbst]” (GMS, AA IV: 414) by the moral law and obtains no other driving-forces, there exists no imperative). So the possibility of rendering respect for the moral law to be the driving-force – “sufficient in and of itself” (RGV, AA VI: 27), as Kant puts it – of the will cannot be designated only by lack of contradiction; in some sense, Kant has to account for what can be termed the ‘real possibility’ of the will to be determined morally. To give a somewhat trivial example: If one sees from her tenth floor window that an old man is about to be run down by a car, it is possible for her to jump out of the window and assist him, for no logical contradiction is involved in such an action. However, we would probably say that she did not have a real possibility – in reference to *our* world and not merely to *a* possible world – of saving this old man. So it is not that this action is not possible per se; rather, what is not possible is *her doing this action* (given the actual world).

As we shall see, in Kant’s case the ‘real possibility’ of carrying out a morally good action is thought of in light of a *specific preliminary background* in which the agent’s will is, by definition, immersed; a moral action, according to Kant, means *overcoming* a non-moral (or not-yet-moral) state of affairs which *already* affects our will. That is: Given that natural driving-forces, which cannot be eradicated by definition,³ are always in play and are, as we shall see, *already* active in some sense, acting morally means a *subordination* of them to the moral driving-force. The agent ought to freely *decide* which of the two kinds of driving-forces – the natural (i.e. ‘self-love’⁴), or the moral (i.e. ‘respect for the moral law’) – would be the subordinated versus the subordinating force.

Here one must be very careful not to confuse *description* with *ground*. The description of a free act (good or evil) entails only two possibilities: Either ‘respect for the moral law’ serves as the subordinated component of the will, and ‘self-love’ as the subordinating (the evil case), or vice versa (morally good). However, the ground for the moral agent to bring about the good or the evil state of affairs is something different; one might choose e.g. to bring about the perverse order – i.e. evil – due to a variety of grounds, e.g. curiosity, boredom, laziness and so on. Put differently, ‘freedom’ (in the sense of ‘freedom of the will’; ‘Willkür’) is not precisely about the subordination relationship between ‘respect’ and ‘self-love’; rather, it is about *our* grounds for bringing about this or that specific order; i.e., it is about “which of the two [‘self-love’

³ See e.g. RGV, AA VI: 36.

⁴ Though Kant distinguishes between different sorts of ‘self-love’ (e.g., ‘mechanic’ and ‘comparative’; RGV, AA VI: 26–7), they all constitute a different *kind* of driving-force, contrary to the moral driving-force. Kant explicitly sees it in this way; see KPV, AA V: 22; RGV, AA VI: 9.

or ‘respect’] he [Er; the free agent] renders the condition of the other” (RGV, AA VI: 36). The word ‘he’ in this sentence plays a crucial role; if it were not about *our* decision concerning the order, but only about the order itself, then a critique in the spirit of Reinhold, according to which the human agent is nothing but a passive observer of the struggle between ‘self-love’ and ‘respect’ would be impossible.⁵ So Kant must not only present us with a description of a morally good action (an action in which ‘respect for the moral law’ subordinates ‘self-love’ and not vice versa); he also needs to show us that the agent has *good grounds* for choosing this – and not the perverse – order of driving-forces.

For Kant, this challenge is particularly difficult; for according to Kant an action which is determined by natural driving-forces (‘self-love’) generates *pleasure*, whereas the moral driving-force (‘respect’) does not. If so, what can serve as a real ground for the agent to subject the pleasurable natural driving-forces to a pure, non-natural moral driving-force whose satisfaction promises her nothing like pleasure? Even if one admits the moral duty, can it be strong enough to make the agent *actually* choose it? One can understand the motivation to subject some natural inclinations to other natural inclinations for the sake of a specific desired purpose. Roughly, this is the logic that guides social contracts à la Hobbes or Rousseau; they will convince the agent that by listening to them she will avoid pain generated by violence, for instance. But Kant’s moral driving-force is something different *toto genere*; it promises nothing and is independent of reward or pleasure. So why should an actual agent subject her natural inclinations to it, even if she is aware of the fact that such a demand exists? In Maimon’s words:

The feeling of respect for the moral law is a motive [=Triebfeder⁶], however not the sole motive, for the determination of the will. Why should I prefer, in cases in which there is a conflict between this motive and other motives, this motive [the moral]? Why should this artificial, factitious feeling [dieses erkünstelte Gefühl] determine my action much more than all other natural feelings? (VND, GW VI: 277)⁷

And this is precisely the problem at stake.

I will begin by concisely describing the problem as Kant saw it and outlining, however generally, the solution he suggested, which is rooted in his disputed concept of the ‘fact of pure reason’ as morality’s point of departure. Then I will

⁵ See Reinhold 2008, 199. For an enlightening discussion of this problem see Noller 2015.

⁶ Maimon explicitly uses ‘Motiv’ and ‘Triebfeder’ as synonyms: “Ich werde in der Folge Motiv statt Triebfeder brauchen” (EGM, GW VII, 462). Kant uses ‘Motiv’ in a similar context as well; see e.g. TP, AA VIII: 282.

⁷ All translations from Kant’s and Maimon’s works are mine.

focus on some of the major aspects of the somewhat forgotten critique of Salomon Maimon in this regard.⁸ I will not argue that Maimon addressed all aspects of Kant's complex solution; in fact, he ignores much of it. However, I believe that ultimately Maimon clearly saw what was at stake and that his pointed arrow is well aimed at the most basic problem which underlies Kant's position. Maimon's attitude is – as far as I can tell – unique in the context of post-Kantian philosophy, in that, as opposed to e.g. Fichte or Schelling, Maimon does not try to trace back the emergence of the categorical imperative to a prior 'deep' metaphysical background.⁹ Rather, for Maimon the 'Sollen' – in the sense of Kant's 'Faktum' – is not original but simply rooted in *an observable natural driving-force* (though a unique one, as we shall see).

2 Why Should We Act Morally according to Kant?

The fact *that* the categorical imperative is (so to speak) an original 'fact of reason' which ought to determine the will is one thing; *why* we should freely render it our main (subordinating) driving-force is another thing. For even if we grant Kant that the categorical imperative presents itself to us as an original 'fact,' it still does not follow that we have any reason to really act according to it. One can acknowledge this fact and still refuse – on good, understandable grounds – to render it the main (subjecting) driving-force of her will. And given that the categorical imperative entails neither threat (if rejected) nor reward (if accepted) – or worse, that it sometimes demands that we should give up our life and our drive for "self-preservation [Selbsterhaltung]" (KpV, AA V: 30) to obey it – it is even reasonable to think that despite acknowledging the demand it imposes upon us we would nevertheless prefer not to obey it. Needless to say: we can also ignore a demand which promises us happiness, for instance; however, in that case such ignorance calls for an explanation, whereas in the moral case our obedience requires an explanation.

⁸ The most detailed study of Maimon's philosophy, which encompasses all aspects of his philosophy, is Kuntze 1912.

⁹ Thus, despite the fact that, as Beiser rightly remarks, Maimon's philosophy opens "a new chapter in the history of post-Kantian philosophy," there is some sense in which, as I will show, it is not correct that "it marks the decisive transition from critical to speculative idealism." See Beiser 1987, 287.

Now it is well known that according to Kant, as I mentioned, the consciousness of the moral law is not just a ‘fact of reason,’ but also an *original* ‘fact of reason’¹⁰. Among other things, this means that a finite understanding can never grasp the way of emergence of this ‘fact.’ In Kant’s words:

How this consciousness of the moral laws is possible or, what counts as the same [einerlei], how freedom is possible, cannot be further explained. (KpV, AA V: 46)

As Kant says time and again, the term ‘original’ in this regard refers to the common understanding.¹¹ This means that the consciousness of this so called ‘fact’ is not unique to philosophers, nor to people who obtain a unique insight, but is accessible to every rational being. Moreover: This ‘fact’ is forced upon us, whether we want it or not, in the same sense – and this is just an analogy – that the outer world forces itself upon our senses (that is, in the same sense in which our cognition has a receptive, not only a spontaneous component). To cite one of Kant’s celebrated examples:¹² If a common agent is commanded to give a false oath which would directly lead to the execution of an innocent man, then she knows – as a matter of ‘fact’ of her consciousness and regardless of the way she will eventually act – that she ought not to do it; just like the philosopher she hears the “voice of reason” (KpV, AA V: 35) and cannot explain its origin. Philosophers after Kant tended not to deny this ‘fact of reason’ or ‘fact of consciousness’ as such (even though, like Hegel for instance, they were discontent with pairing the term ‘fact’ with ‘reason’), but rather its originality.¹³ That is: they did not suggest a new morality, but a new grounding of the same morality (Maimon, for instance, explicitly speaks not of a new principle of morality, distinct from the Kantian, but rather of a new “formula” [VND, GW VI: 275]¹⁴ of the same principle).

As I mentioned earlier, it is not enough to argue that morality is rooted in an original ‘fact of reason’; for, since this fact presents us with a demand, it should entail the possibility of its *actual* realization. Kant ought to further show us what can serve as the ground for the common, real agent to prefer to subject her other, pleasurable natural driving-force, rooted in ‘self-love,’ to the moral driving-force. Moreover: according to Kant, if the finite agent renders respect for the moral law

¹⁰ For a discussion of some subtle aspects of this issue see Wolff 2009; Ware 2014.

¹¹ Sometimes he even refers to it as “the *most* common understanding [der gemeinste Menschenvernunft] “; see TP, AA VIII: 286; my emphasis.

¹² Cf. KpV, AA V: 30.

¹³ See e.g. Fichte 1995, 13f.

¹⁴ See Zac 1988, 224ff.

her driving-force, sufficient by itself, of her will – this involves what Kant terms “breaking [Abbruch] with all inclinations” (KpV, AA V: 72) – then not only is pleasure not the outcome of such an act, but on the contrary: This move is accompanied by pain and even by “humiliation [Demütigung]” (KpV, AA V: 75) according to Kant. In light of this, what can genuinely serve as a ground to act morally, i.e. to subordinate ‘self-love’ to ‘respect’?

Kant well understood the challenge he was facing. He knew that the moral ‘driving-force’ should incite real – meaning, *finite* – agents to action. It is easy for a God to act morally; a God cannot act otherwise,¹⁵ since, as we have seen, God’s will is determined only by the moral law.¹⁶ This challenge threatens to render Kant’s whole moral project inapplicable; for Kant cannot simply add a certain reward that would motivate the moral agent to act morally, for this would be a ‘contradictio in adjecto’ in Kant’s terms; even the expectation of a reward renders the act not purely moral.

So what is Kant’s way out of this maze? It pertains precisely to the finitude of human will. As Kant famously determines,

[w]ithout any reference to a purpose, no determination of the will can take place in a human being at all. (RGV, AA VI: 4)

Kant excludes from the moral account only a concept of an end which *precedes* the will,¹⁷ and not every concept of an end as such; this a-priori end, which does not contradict morality but rather enables its applicability, is called the ‘Highest-Good.’ This is why Kant denotes:

If [...] the Highest-Good [the ‘end’ which does not precede the will, “the object of the power of desire of rational finite beings,” KpV, AA V: 110] is impossible according to practical rules, then the moral law which commands us to further this good must also be fantastic and aimed at empty imaginary purposes, and hence in itself false. (KpV, AA V: 114)¹⁸

So this is why the ‘Highest-Good’ must be thought of as possible; without its possibility, morality simply cannot be carried out by finite beings. Moreover: Kant is in need of a concept of an end in order to render morality really possible given that the human will’s point of departure is not neutral; it is *already* ‘sunk’

¹⁵ See Insole 2013, 86.

¹⁶ At least as far as actions ‘in’ the world are concerned, but not necessarily concerning the ‘beginning of the world’; see Prolegomena, AA IV: 344Fn. I cannot dwell further on this point here.

¹⁷ See e.g. RGV, AA VI: 4.

¹⁸ See also MS, AA VI: 385

(though not yet fully determined – for in that case there is no room for freedom) in other, material ends:

Given that the sensible inclinations mislead and attempt [verleiten] for ends (as the material of choice) that can be contrary to duty, lawgiving reason cannot resist and refuse [wehren] their influence without giving a moral end which stands against them, i.e. without suggesting an a-priori end, which is independent of inclinations. (MS, AA VI: 380–1)

So a real possibility to morally determine the will must show us first – with the aid of a conception of an a priori end – how to pull the will out of material ends in which we are already ‘sunk-in.’ ‘Morality’ can serve as a real driving-force for finite beings only through the mediation of the representation of this a priori end.

Now Kant was well aware of the fact that if this representation of an a priori end serves as the ground for a moral action, then one can argue against him that at the end of the day this ground entails a representation of the agent’s *own happiness*; for the ‘Highest-Good’ is, after all, a representation of a moral world, in which there is a correlation between virtue and happiness. Such a driving-force would render morality impossible, for it is not pure; carrying out the ‘Highest-Good’ would be a mask under which an egoistic driving-force (the agent’s own happiness) is hidden. Kant’s rejoinder is complicated, and I cannot dwell on it in detail here. Let me just quote one relevant aspect of it:

[A]lthough the concept of the highest-good, as that of a whole in which the greatest happiness is presented as linked in the most exact proportion with the greatest degree of moral perfection (possible in creatures), includes also my *own happiness*, yet the determining basis of the will that is instructed to further the highest-good is not this happiness but the moral law. (KpV, AA V: 129–30)

So Kant does not deny that the agent’s own happiness is included in the representation of the ‘Highest-Good’ – the ultimate ground of a moral action; however, he only denies that this happiness, despite it being included in the representation of the ‘Highest-Good,’ is that which actually determines her will. I will not try to defend or to attack Kant’s position here, though it is interesting to see whether it can be defended, psychologically as well as conceptually; I brought it up only to demonstrate that no one can blame Kant for ignoring the difficulty at stake.

This is not the end of it; Kant presents us with an argument in which the conclusion is that the point of departure of every¹⁹ finite being is evil;²⁰ in light of

19 Cf. RGV, AA VI: 21, 25, 29, 32. For Kant’s alleged ‘proof’ of this claim see Buchheim 2001.

this, even if one chooses the good, other people would tempt her time and again to do evil.²¹ Thus, attaining the ‘Highest-Good’ requires that people unite and establish a church, i.e. “a joint ethical entity” or “society of laws of virtue [Tugendgesetzen]” (RGV, AA VI: 94). Without getting into every minute detail of Kant’s rich account in this regard,²² you can see how Kant was troubled by the question regarding the possibility to carry out morality, given the *real* background of human agents, and how many mediated links it takes to show that this move is actually possible.

But what if a pure driving-force à la Kant, i.e. one that is categorically distinct from natural driving-forces, cannot *really* – not logically, for no contradiction is involved – determine the will? This was precisely the question Maimon raised. It is well known that concerning the realm of theoretical philosophy, Maimon raised doubts about whether Kant’s synthetic ‘a priori’ principles actually apply to sense intuitions (he decisively answers: No).²³ Concerning Kant’s moral philosophy, Maimon raises a similar objection: No fact seems to correspond to Kant’s pure principle. So the problem of *quid facti* is brought to the fore by Maimon in the moral realm as well; and this, as we shall now see, has something to do with Kant’s disputed point of departure concerning morality, i.e. with the alleged *originally* of the ‘fact of reason.’

3 Maimon’s alternative: A Preliminary Reflection

Before delving into Maimon’s alternative to morality’s point of departure – that is, an alternative to the ‘fact of consciousness’ from which Kant departs – it is of interest to take a brief look at one of Maimon’s sober reflections on the very ‘pureness’ of the transcendental project as such.

According to Maimon, there are some terms in the language which become ambiguous as time passes and knowledge increases; this holds true also for concepts in philosophy such as ‘pure,’ ‘a priori’ and ‘empirical’²⁴. Maimon adds:

It is not to deny that the critical philosophy gave a more definite meaning to these terms than before. However, it only gave it to the extremes [...] no attention was given to the mid-

²⁰ Cf. RGV, AA VI: 94.

²¹ Cf. RGV, AA VI: 94.

²² For an account see Kravitz 2018.

²³ For an illuminating presentation of Maimon’s theoretical philosophy as a whole see Buzaglo 2002.

²⁴ Cf. EGM, GW VII: 460.

dle stages; but there are such important kinds of cognitions [Erkenntnißarten] in these middle stages! (EGM, GW VII: 460)

According to Maimon, even Kant's metaphysics of nature has empirical representations at its foundation, and this holds true for the 'a priori' of the transcendental philosophy as well. Now it is not the case that Maimon rejects the concept of 'a priori' as such; rather, he accepts it to the extent the "nature of things" (EGM, GW VII: 461) permits it. And in the case that 'a priori' knowledge is not sufficient for the determination of the object at stake, that is: when the 'a priori' seems not to correspond to the special nature of things in a specific realm, it is sometimes advisable according to Maimon to let go of what he occasionally terms the "strictness [Strenge]" of its demand and to see whether an alternative way, which might achieve the same end, is possible. This insight has a special meaning with regard to the moral realm, where, as Maimon puts it, the aim is not to satisfy our desire for knowledge, but making us morally better, i.e. *truly* moving us to action. The difficulties Kant faces regarding this point are symptomatic in a way, mainly because Kant insisted on grounding morality on a special (pure) 'fact,' which is categorically distinct from natural driving-forces which actually drive people to action. By doing so, Kant seems to have lost track regarding the aim of morality and the possibility of its application at the cost of retaining the purity of the principle.

I think the fact that Kant claims that his morality retains its validity even if not a single agent has ever acted in a Kantian moral way thus far²⁵ is a strong indication of this problem. I stress the term 'indication' here; it is by no means a 'proof' which invalidates the foundations of Kant's moral theory. For Kant justly argues that this fact does not render the moral demand impossible; it can be seen merely as a contingent – however interesting – fact about human history. However, it may also cast some serious doubts on the applicability of morality, whose aim is to move "world-beings [Weltwesen]" (KU, AA V: 447) – i.e. finite beings and not a God – to action. Put differently, this contingent fact might attest after all to an essential problem; it might tell us that Kant's pure project went too far away from real human beings and the driving-forces that they might really choose to determine their will. It is as if Kant desperately tries to bridge the gap between his understandable inspiration to ground morality once and for all on pure grounds and real agents; for having shown *what* a moral action means (choosing 'respect' as the subordinating component of the will) still does not amount to showing *that* real agents have a sufficient, reason-

25 See, for instance, GMS, AA IV: 408; PT, AA VIII: 384–5.

able ground to choose it, given that it generates no pleasure and even humiliates.

This is the reason why Maimon's moral point of departure – the so-called 'fact' – does *not* pertain to a 'pure voice of reason':

I see myself compelled by certain *Facta*, which are not based on pure reason but on general observation [allgemeine Beobachtung] of human nature. For reason itself is nothing but the observed Faktum of human nature! (EGM, GW VII: 461)

Maimon's modest concept of 'observation' is the key, which ought to capture the "nature of things" mentioned above; 'observation' means, first and foremost, an immediate affinity with real agents and their real possibilities.

4 An Alternative 'Faktum' As Morality's Point Of Departure

Maimon has a handful of reasons to reject Kant's original 'fact of reason' as morality's point of departure. In fact, he dedicated his last text – *Der moralische Skeptiker* – to develop an insightful, highly sophisticated attack on Kant in this regard. I have written about it in detail elsewhere²⁶, and I cannot reconstruct Maimon's complex argument here. Let me just say a few relevant words about it.

Maimon points his arrow not at Kant's 'fact of reason' as such, but rather at the originality Kant ascribes to it. The principle which guides Maimon's considerations is the following: "It stands in contrast to any scientific method to assume a distinct, special principle [ein eignes Prinzip] in order to explain an appearance [Erscheinung], if this appearance might be explained by principles that are already known" (MSK, GW VII: 542f). Notice two things which are implied here: (a) Maimon's position is not moral skepticism in the sense of denying the absolute moral demand which is encapsulated in the 'fact'; Maimon accepts the absoluteness of duty, as it appears in Kant's 'fact,' and merely adds that one can further try to suggest ways to unveil its way of emergence; (b) Maimon does not argue that Kant's claim that the 'fact' is original is wrong; rather, he uses a conditional: *If* we could explain the existence of the same 'fact' in our consciousness with the aid of already known principles, *then* we would have no reason to accept Kant's claim that it is a *sui generis* 'fact.' Moreover: Even if we have not *yet* discovered how an action, which common reason thinks was determined by

26 See Kravitz 2017.

duty, can be also explained by “other [natural] drives [Triebe] and inclinations [Neigungen]” (VND, GW VI: 289), we cannot conclude that “such an account cannot be given” (VND, GW VI: 289). This means that according to what Maimon terms “the scientific method,” explaining an appearance using known principles is a *better* explanation, but not necessarily the *right* explanation. From a meta-physical point of view, Kant’s claim might be the right one. However, since we cannot penetrate the supersensible, we might have methodological grounds to pragmatically *prefer* the known explanation, given that we can actually suggest one.

The reason for Maimon to suggest an explanation which is based on already established principles is also rooted in his essential suspicion of the judgments of ‘common reason’ (at some point Kant writes: “[E]ven a child at the age of nine years” [TP, AA VIII: 286] can understand our duty). In Maimon’s words, “this appealing to the common reason [gemeiner Menschenverstand] is a highly unfortunate [misslich] thing” (EGM, GW VII: 453). For what the “most common human reason” (TP, AA VIII: 286) judges as original can turn out to be, after all, a result of some psychological illusion.²⁷ We ought to begin in a different way, one that cannot succumb to possible deceptions:

Kant develops the principle of morality in light of the concepts of common, ordinary reason. I, on the other hand, according to my skeptical method, suspect common reason insofar as one can explain its concepts through psychological deception. This is why, in order to ground the moral principle, I must find a new way in which, as well, a fact serves as its basis, but in which no psychological deception can take place. (VND, GW VI: 276)

So if duty, which the common understanding judges as an *original* “voice of reason,” could be explained as a *consequence* of a natural driving-force – something which we can really observe in experience – then we would have no reason to prefer the Kantian explanation. We abandon, then – for the sake of proving the possible intelligibility of the “scientific” alternative – the description of the ‘fact’ as a judgment of common reason concerning the originality (i.e. something not of nature) which appears in our consciousness, and try to see if this ‘fact’ might be also explained as a consciousness of a (natural) drive that can be observed. In a way, Maimon turns Kant on his head: Kant first sees to it that the *possibility* of a moral act is secured, and then seeks, as we have seen, the conditions of its *actual realization*. Maimon explicitly reverses the order; he first observes a *real* (natural) drive which without any doubt *already* occasionally determines our will, and asks whether it is *possible* to elicit morality from it. If it turns out that morality,

27 See, for instance, MSK, GW VII: 549; EGM, GW VII: 453; VND, GW VI: 276.

indeed, can be deduced based on this already actual drive, then morality's reality, as Maimon says, secures the possibility.²⁸ The precedence of the 'quid facti' makes its appearance in the moral realm as well.

Two issues are, thus, at stake. The first concerns the character of Maimon's alternative – real, observable – 'fact,' and the second concerns the possibility to deduce the principle of morality from it. Here are two of Maimon's revealing formulations in this regard, which encapsulates, as it were, Maimon's way of tackling these two issues:

I am aware of a drive [Trieb] to the higher perfection [Vollkommenheit] (to the highest possible development of my powers). That is, I obtain a drive to develop my cognitive capacity [Erkenntnisvermögen] [...] i.e. a drive to cognize the truth [Erkenntnis der Wahrheit]. The character of truth, however, is universal validity [Allgemeingültigkeit]; thus, I obtain a drive of rendering my representation [Vorstellungen] to be universally valid [Allgemeingültigmachung]. When I make this drive universal, and apply it not only to my cognitive capacity but also to my will, then one can explain on this basis and without any deceptions the judgment of the common reason on the morality of actions [...]. [F]rom [the drive for the cognition of truth] I deduce [herleiten] the moral law and the possibility to act according to it. (VND, GW VI: 278)

Two [...] sentences are at stake here: 1) "a being gifted with reason [ein vernünftiges Wesen] determines itself in a necessary way to think according to the laws of truth [this is Maimon's point of departure]"; 2) "a being gifted with reason determines itself in a necessary way to act according to the demands of duty" [this is Kant's point of departure]. The first sentence is given to us as a fact of consciousness [Faktum des Bewusstseins]; its possibility is demonstrated [or shown, sustained; *dargetan*] through its reality [Wirklichkeit] in a way that cannot be doubted. The second sentence, however, is given to us likewise as a fact of consciousness in itself; its possibility, however, cannot be demonstrated through its real use [Gebrauch], for its use can always be doubted. I thus examine whether these two predicates – truth and duty – share a common feature [Merkmal], which is the actual [das eigentliche] predicate of the subject – of the beings gifted with reason. (VND, GW VI: 290f) In order now to explain this, we render the first fact to be general [allgemeine], and assume that we obtain not only the drive of rendering our cognition universal [Allgemeinmachung unserer Erkenntniß], but also the drive of rendering our will universal [Allgemeinmachung unsers Willens]. The second fact, namely the moral law, can be explained on this ground. (VND, GW VI: 313)

As for the first issue, then, Maimon's alternative 'fact' refers to the observable, natural drive to achieve perfection regarding knowledge (or simply: to know the truth; Maimon does not refer to knowing the truth of a *certain matter*, but to *generally* quest for truth). And as for the second issue, Maimon's method of deducing morality from this fact – i.e., to elicit morality's possibility from the

28 Cf. VND, GW VI: 290.

actuality of the fact – is carried out by means of a mediated concept: ‘universal validity.’ For given that (i) we truly possess a drive to develop our cognitive capacity, i.e. to render our representation universally valid (and we can observe that people are really determined by this drive, just as they are driven by e.g. sensual drives), and assuming that (ii) this observable drive can be applied to our will as well, and not only to our cognition (for we actually want our will to be universal), then (iii) we can try to deduce the moral law from it by grounding it in the concept of ‘universal validity.’ For ‘general validity’ can surely serve as a common feature here:

Duty [Pflicht] is the representation [Vorstellungen] of the necessary restriction [Einschränkung] of the will of a being gifted with reason through the possible will of other beings gifted with reason as well. It follows that truth and duty share an essential common feature: universal validity. I cannot know and recognize something to be true [für wahr erkennen] which cannot be valid as true for any other being gifted with reason. [And] I cannot want something out of duty if it cannot be wanted by any other being gifted with reason. (VND, GW VI: 291f)

Let us take a more precise look at the differences between Maimon and Kant in light of all this.

5 On Pleasure and on the Essential Precedence of Obedience

First and foremost, it is important to notice that according to Maimon this “undoubted fact” (EGM, GW VII: 454) is essentially applied to “*every* natural being” (EGM, GW VII: 454; my emphasis) and not only to human beings. For every natural being can be described with certainty as determined through what Maimon terms “born drives [angeborene Triebe]” to achieve *its* perfection, i.e. “to render real [Wirklichmachung] what is possible for him” (EGM, GW VII: 454). It follows that for human beings, just like for other natural beings, a satisfaction of this drive – just like with regard to every other drive – necessarily involves pleasure.²⁹ It does not mean that the *kind* of pleasure elicited from a possible satisfaction of this drive is identical to the pleasure generated by satisfaction of sensual drives – Maimon points out several essential differences here,³⁰ but what is crucial is the

²⁹ For a concise discussion of this point see Baumgardt 1963, 203f.

³⁰ For instance, the drive for knowing the truth, as such, and despite it being *of* nature, is distinct from *all other* natural drives alike; it is an ‘end-in-itself’ or ‘sublime’ (EGM, GW VII: 455), for

fact that pleasure – regardless of its exact kind – comes to play a substantial role here. This is why Maimon thinks that actual agents would obtain – according to this description – a real ground for choosing morality. For contrary to Kant, who tries to show that actual agents would nevertheless choose to act morally even if this would generate humiliation for them, Maimon does not face a similar challenge; if the moral drive – at least as far as the factuality of generated pleasure is concerned – is not different *toto genere* from natural drives, then a much more solid ground to act morally is given. So Maimon does not find himself entangled in a (in his eyes) desperate attempt to link a *pure* driving-force to ‘Weltwesen.’ And only through the moral law (that is, by choosing to obey it) can this unique drive be satisfied.³¹

But how does Maimon’s morality – i.e. this restriction of one’s will through the possible will of other rational beings, a restriction without which the universalization [Allgemeinmachung] of one’s will and the pleasure it brings with it cannot take place – actually work? At first glance, Maimon’s principle of morality seems strikingly similar to Kant’s:

Act in a way in which your will (through which your actions are determined) could be thought of every time as the will of every other being gifted with reason [...]. (VND, GW VI: 309)

However, Maimon’s deviation from Kant’s intent is, according to him, fundamental, and has something to do with the fact that in the moral realm ‘universal validity’ conditions ‘objectivity,’ i.e. it has something to do with Maimon’s claim for the precedence of the factual existence of a *plurality of subjects* to the *objective form* of the moral law:

The principle of universal validity assumes [voraussetzt] all beings gifted with reason are in mutual obligations [wechselseitige Verpflichtung] towards each other. (VND, GW VI: 312) Truth as determined by theoretical reason is in itself [an sich] an objective, and through it universally valid. In practical reason it is the opposite; the will which is determined by it is universally valid, and only through it objective. Through the first [theoretical] I think: A is B; here I abstract from every connection [Beziehung] to myself and consider only the relation [Verhältnis] of the given objects A and B with each other [...] [in mor-

it is a drive to determine the will against *all* other sensual drives and inclinations. In addition, other drives are always related to a given external object, whereas the drive for knowing the truth is unrelated to an object as such [Objekt überhaupt] (EGM, GW VII: 468). Further, it is the only *disinterested* drive (VND, GW VI: 279), not because it is not in our interest to satisfy it – as every natural drive, it calls for a satisfaction as well – but rather due to the fact that it is unrelated to satisfaction from a *single* object.

31 Cf. EGM, GW VII: 455.

ality, however] I first confirm [versichern] the universal validity of the action (i.e., I confirm that it does not stand against [zuwider] the purpose [Zweck] of the rational subject), before I convince myself that the action is objective (that the will was not determined by my individual subject). That is: In order for me to know that no individual determining grounds were entailed in my will, I must first know that my will can subsist [bestehen kann] with every other individual determination. [...]. Through theoretical reason the human being is an absolute law-giver, a monarch in the spirit of Hobbes, who has the power to legislate laws which are universally valid without the need to investigate first whether these laws stand in contradiction to the will of some member of the society. [...]. [T]hrough practical reason, on the other hand, the *human agent must first learn to obey before wanting to govern through his law-giving*. That is, the human agent [in the practical realm] must first determine the negative cases in which her will (because it stands against the will of some other rational being) cannot be valid as a law, before she becomes a positive law-giver. The human being is a member of an accomplished republic [Mitglied einer vollkommenen Republik]. (VND, GW VI: 315–317; my emphasis)

Theoretical philosophy, then, is about ‘objectivity’ in the pure sense of the word; *if* something is objectively true (=‘objectivity’), *then* and on this ground it has validity to all other subjects (=‘universal validity’). In the practical realm it is precisely the opposite: There, I first take into consideration the conjectural will of others (=the ‘universal’ aspect); thus, the precedence of what Maimon terms ‘the republic’ (I will say something about its precise character in due course), i.e. the concept of ‘obedience,’ is brought to the fore. The very notion of ‘moral possibility’ is intelligible only with reference to the preceding, factual existence of other (rational) subjects. Thus, whereas Maimon argues that Kant thinks that “I must be able to will [ich muß wollen können] that every other being gifted with reason would also enact [machen soll] the maxim that I make for myself” (VND, GW VI: 309), Maimon holds that “I must want the action which every other being gifted with reason is able to want me to exercise.” (VND, GW VI: 309)³² According to Maimon, then, the so-called ‘republic’s’ expectation that I would choose an action stands at the center.

In order to better grasp the difference Maimon is trying to point out, it is useful to examine one or two concrete examples. Take e.g. the problem of suicide. According to Kant and without getting into some subtle distinctions he makes in this regard,³³ suicide is forbidden in light of obligations of the moral agent to-

³² Maimon’s precise formulation is not easy to translate here: “ [...] ich [muß] diejenige Handlung wollen, von der ein jedes vernünftiges Wesen wollen kann, daß ich sie ausüben soll.”

³³ See e.g. Wittwer 2001.

wards herself [Pflichten gegen sich selbst]³⁴; even if you are the last person on earth, suicide is still an act against one's duty. Thus, it is independent of the factual being of others, and has to do only with the *form* of the law, which absolutely forbids it. It seems that in light of the character of 'obedience' mentioned above which marks the moral realm according to Maimon, no room is left for the very conception of 'Pflichten gegen sich selbst.' And indeed, this is precisely Maimon's line of thinking: *Given* that duty is, as mentioned above, "the necessary restriction [Einschränkung] of the will of a being gifted with reason through the possible will of *other beings* gifted with reason as well," i.e. given that duty makes no sense without the context of the factual existence of other rational beings whom one ought to obey, then "there is, in point of fact, no duties towards oneself" (VND, GW VI: 311).

However, this does not mean that Maimon's morality enables suicide; as already mentioned, Maimon wants to account for the emergence of the *same* 'fact,' i.e. to ground – though in a different, "scientific" way – essentially the very same morality. Rather, according to him, "from the principle: My will must be universally valid as the will of a rational being, it follows that while the social [gesellschaftlich] (real or only possible) connection of rational beings [...] must be their general will, I am obliged to the rest of the members of the community to [...] retain my life" (VND, GW VI: 312). Obligations against oneself, then, must be explained in terms of obligations towards others; hypothetically, if a single person exists alone – say, a Robinson Crusoe of the whole world (and not only of his island) – then he obtains no duty not to commit suicide.

Another example: Say that the following maxim guides me: "To use all safe means to enlarge my fortune" (VND, GW VI: 309). Now say that – as in one of Kant's celebrated examples³⁵ – someone deposited money in my hands, died shortly afterwards and did not leave any document saying that she gave me money as a deposit. According to Kant, as Maimon understands him, I ought to ask myself if the maxim can have the form of a law, according to which it is permissible for everyone in a similar situation to deny that money was deposited if it cannot be proven that this deposit took place. Maimon wished to ground the same result – that one ought to give the deposit back – in a different way; "I ask myself whether every rational being (and thus also the person who gave me the money) can want me to act according to this maxim (denying the acceptance of the deposit)" (VND, GW VI: 310). This is impossible according to Maimon not

³⁴ For Kant's position see e.g. MS, AA VI: 422: "Die Selbstentleibung ist ein Verbrechen (Mord). Dieses kann nun zwar auch als Übertretung seiner Pflicht gegen andere Menschen [...] betrachtet werden; – aber hier ist nur die Rede von Verletzung einer Pflicht gegen sich selbst [...]."

³⁵ Cf. TP, AA VIII: 286f.

because – in a Kantian spirit – it contradicts the form of the lawfulness [Gesetzmäßigkeit], but due to the fact that *the owner of the deposit herself* cannot want me – when giving me the money – to deny it if she were to suddenly die. Maimon's attempt at rooting morality in a concrete, realistic background makes its appearance again.

However – and with this remark I conclude – there is a strong sense in which Kant was, as far as I can tell, much more 'realistic' so to speak than Maimon in the moral realm. For it seems to me clear that Maimon refers to a very narrow sense of 'republic,' that is, of 'community' in the moral context: It is a universal community of agents, whose *reason* is the basis which connects them.³⁶ This is parallel, in a way, to Kant's intent in his celebrated stipulation that "morality leads inescapably [unumgänglich] to religion" (RGV, AA VI: 4). For Kant thought, as well – though as a conclusion, not as a point of departure like Maimon – that in order to realize morality (in the sense of rendering it real) one must "inescapably" ground a "republic governed by laws of virtue [Republik nach Tugendgesetzen]" (RGV, AA VI: 98). If Kant would have stopped here, then one would be justified in holding that both Kant and Maimon saw the inescapability of the concept of a 'republic' of rational agents in the moral realm, but whereas for Kant it served as an "inescapable" means to realize morality as a pure project, for Maimon it served as morality's point of departure. However, Kant also adds – and here I can only mention his conclusion, not his whole intricate argument – that due to some "unique weakness of human nature" (RGV, AA VI: 103) this 'pure (wholly rational) republic' cannot *really* influence people and drive them to action if it is grounded on reason *alone* (it cannot be achieved "through bare reason alone" [nicht durch bloße Vernunft])" (RGV, AA VI: 105)³⁷. Some additional, contingent component, in which the 'republic' would appear in the form of a positive 'revelation,' of a concrete, contingent story which is never wholly identical to reason alone,³⁸ must be added. Kant held, thus, a much 'thicker' conception of the *real background* needed to drive actual agents to action. This, however, deserves a discussion of its own.

³⁶ This does not mean that Maimon did not judge positive religions – first and foremost Judaism and Christianity – from a philosophical point of view; this, however, is a different issue. See Buzaglo 2002b.

³⁷ See also SF, AA VIII: 49.

³⁸ For "every historical belief entails by definition – for it is not identical to pure reason – the consciousness of its contingency [Bewusstsein seiner Zufälligkeit]"; RGV, AA VI: 115.

Bibliography and Abbreviations

Kant's works are cited according to *Akademie-Ausgabe* [AA], Berlin 1900f.

GMS	Grundlegung der Metaphysik der Sitten
KPV	Kritik der praktischen Vernunft
RGV	Die Religion innerhalb der Grenzen der bloßen Vernunft
TP	Über den Gemeinspruch: Das mag in der Theorie richtig sein, taugt aber nicht für die Praxis
Prolegomena	Prolegomena zu einer jeden künftigen Metaphysik, die als Wissenschaft wird auftreten können
MS	Metaphysik der Sitten
KU	Kritik der Urteilskraft

Maimon's works are cited according to *Gesammelte Werke* [GW], ed. Valerio Verra, Hildesheim 1965 ff.

EGM	Ueber die ersten Gründe der Moral
MSK	Der moralische Skeptiker
VND	Versuch einer neuen Darstellung des Moralprinzips und Deduktion seiner Realität

Baumgardt, David (1963): "The Ethics of Salomon Maimon (1753–1800)." In: *Journal of the History of Philosophy* 1, 1963, pp. 199–210.

Beiser, Frederick C. (1897): *The Fate of Reason; German Philosophy from Kant to Fichte*. Cambridge: Harvard University Press.

Buchheim, Thomas (2001): "Die Universalität des Bösen nach Kants Religionsschrift". In: V. Gerhardt, R. Horstmann und R. Schumacher (Eds.): *Kant und die Berliner Aufklärung*. Berlin, New York: De Gruyter, pp. 652–661.

Buzaglo, Meir (2002): *Solomon Maimon: Monism, Skepticism and Mathematics*. Pittsburgh: University of Pittsburgh Press.

Buzaglo, Meir (2002b): "Salomon Maimon: Christianity, Judaism and non-Euclidean Geometry." In: Christoph Schmidt, Bernhard Greiner (Eds.): *Arche Noah: Die Idee der Kultur im deutsch-jüdischen Diskurs*. Freiburg: Rombach Verlag.

Fichte, Johan Gottlieb (1995): *System der Sittenlehre*. Hamburg: Meiner.

Insole, Christopher J. (2013): *Kant and the Creation of Freedom. A Theological Problem*. Oxford: Oxford University Press.

Kravitz, Amit (2017): "Von der Unmöglichkeit der Kant'schen Freiheitslehre nach Salomon Maimon". In: Sasa Josifovic, Jörg Noller (Eds.): *Freiheit nach Kant. Tradition, Rezeption, Transformation, Aktualität*. Leiden/Boston: Brill, pp. 187–207.

Kravitz, Amit (2018): "Innerhalb der Zeit, außerhalb der Geschichte. Zu Kants Auseinandersetzung mit dem Judentum in der *Religionsschrift*." In: Amit Kravitz, Jörg Noller (Eds.): *Der Begriff des Judentums in der klassischen deutschen Philosophie*. Tübingen: Mohr Siebeck, pp. 25–43.

Kuntze, Friedrich (1912): *Die Philosophie Salomon Maimons*. Heidelberg: Carl Winter.

Noller, Jörg (2015): *Die Bestimmung des Willens. Zum Problem individueller Freiheit im Ausgang von Kant*. Freiburg/München: Karl Alber Verlag.

- Reinhold, Karl Leonard (2008): *Briefe über die Kantische Philosophie. Zweyter Band*. Basel: Schwabe.
- Ware, Owen (2014): "Rethinking Kant's Fact of Reason". In: *Philosopher's Imprint* 14, pp. 1–20.
- Wittwer, Héctor (2001): "Kant's Ban on Self-Killing (Suicide)." In: *Kant-Studien* 92, pp. 180–209.
- Wolff, Michael (2009): "Warum ein Faktum der Vernunft ein Faktum ist. Auflösung einiger Verständnisschwierigkeiten in Kants Grundlegung der Moral". In: *Deutsche Zeitschrift für Philosophie* 57, pp. 511–49.
- Zac, Sylvian (1988): *Salomon Maïmon. Critique de Kant*. Paris: Ed. du cerf, 1988).

Tom Giesbers

Drive as a Constitutive Element of Practical Action in Jacobi and Fichte

Abstract: This chapter offers an analysis of the argumentative and semantic usage of the word “Trieb” (drive) as it is offered by Johann Gottlieb Fichte, in the lectures on morality that he presented in 1796. Fichte offers an account of drive that is essential to his philosophical perspective and is intimately tied up with freedom and the will. After demonstrating this systematic complexity of the term, its similarity to Friedrich Heinrich Jacobi’s less argumentatively comprehensive notion of drive will be examined, as well as to what degree Fichte is indebted to Jacobi’s texts in his adoption of this notion. Finally, it will be shown how the shared notion of drive, as a distinctly different semantic context of the word, offers a greater applicability and inclusivity than earlier distinct usages of the word, such as the volitional, zoological, botanical and literary uses. Combined with Jacobi and Fichte’s wide-ranging impact on German publications during this period, this explains why this period boasts the greatest increase in the use of the word “Trieb” that has ever occurred, paving the way to its complex adoption by various philosophers and psychologists during the 19th century.

1 Varieties of “Trieb”

A Google NGram graph of the use of the word “Trieb” in German publications¹ shows several notable increases in the usage of the word, based on the amount of available digitized publications available to the database. Obviously, this is not entirely accurate, since it is limited to the material available on Google Books, which is still a significant amount. The available data is extensive enough to be representative of trends in word usage. Jacobi’s and Fichte’s use of the term can be located during the sharpest spike in usage of the term in the three-hundred years that it measures. The period of the largest rise in the graph roughly corresponds to the rise of Kantianism and German idealism, starting more or less in 1783 and ending in 1801. The graph shows that the word was in use before that time, but we can surmise that something significant changed in its semantic usage which allowed it to be amenable to such intense usage in print. The hy-

¹ https://books.google.com/ngrams/graph?content=Trieb&year_start=1700&year_end=2000&corpus=20&smoothing=3&share=&direct_url=t1%3B%2CTrieb%3B%2Cc0 (12–31–2018).

pothesis of this chapter is that this change is due to the transformation of the term by Jacobi and Fichte, who were widely read and influential authors during this period. In order to illustrate the tenability of this hypothesis, I will examine the exact nature of their transformation of the term. In the process, I will show what kinds of applications of the term Jacobi and Fichte introduced, which will in turn provide an explanation of the popularity of the term during this period.

In order to understand the semantic landscape of “Trieb” that Jacobi and Fichte were operating in it is important to note how varied the reference of the word actually was before they transformed it. Grimm’s *Deutsches Wörterbuch* distinguishes several of these senses by citing various authors who used the word in the 18th century. Today, perhaps the most familiar of these is the botanical reference, in the sense of “sprouting” in English, of buds or pods in plants. The *Deutsches Wörterbuch* also observes a literary sense of “a striving that is not aimed towards a determined object” (*Wörterbuch* 1952, 436). Yet another sense that is observed is “a not precisely defined inner power, which makes itself felt” (*Wörterbuch* 1952, 446). This latter sense can perhaps best be described as volitional drive, due to its highly personal nature. Finally, there is a fourth sense: a “Trieb” which is used more or less synonymously with “Instinkt” by authors who published on the philosophy of mind and anthropology, like Christian August Crusius, Ernst Platner, Johann Georg Heinrich Feder, Hermann Samuel Reimarus and Moses Mendelssohn. The relevant published works of these authors can be located in the rise in the graph prior to the one we are currently concerned with, and were likely influenced by zoological accounts.² Their use was probably the most recent transformation of the term, although it was not nearly as widely used as the transformation that we are concerned with.

This gives us the following scope of the earlier semantic use of “Trieb”:

- i. Botanical – Sprouting, buds, pods.
- ii. Literary – Striving towards an undetermined object.
- iii. Volitional – Undefined inner power.
- iv. Zoological – Animal instinct.

Although there are undoubtedly more senses to be found, these four senses were the most prevalent at the time in which Jacobi and Fichte transformed the term. It should be observed that, although it might seem that there is an obvious conceptual overlap between these senses, the botanical, literary, volitional and zoological contexts were, on the whole, unrelated and strictly separated by their

² Although it is a fairly late example, see, for instance, Georg Forster’s references to “Trieb” (Forster 1783, 128, 192).

contexts. Several authors (such as Goethe and Herder) used two or more of these senses, depending on the meaning that they sought to convey and the *type* of text they were writing.³ This tells us that it was reasonably clear to these authors that the word had different meanings in different contexts.

In what way did Jacobi and Fichte deviate from these established contextual distinctions? Essentially, they found a way to effectively incorporate several or all of these separate senses within a comprehensive *philosophical* account. It might even be the case that our intuition that these four senses overlap is the direct result of the fact that it has become, in the wake of Jacobi's and Fichte's innovations, an established practice to have these senses of "Trieb" overlap. If this is the case, the introduction of Jacobi's and Fichte's integrative philosophical notion of "Trieb" resulted in a semantic shift towards a notion of drive which has retained much of this complexity throughout its transmission to Arthur Schopenhauer, Sigmund Freud, and to our own time. In fact, it is only with the advent of this sense of "Trieb" that it makes sense to translate it as drive, due to its history of being translated as such in Freud's texts.

What exactly did this transformation of the notion entail? In what follows it will become clear why the notion of drive became so intimately entangled with the problems of the *will* and *freedom*. Fichte played the largest role in popularizing the notion among his students, and in his work we can see the clearest shift from a Reinholdian volitional sense of drive towards a Jacobian philosophical sense of drive. This will facilitate the isolation of Jacobi's exact innovation and will make it possible to locate the relevant passages in his work through Fichte's most likely reading of Jacobi's publications. The advantage of this approach is that Jacobi's work, which is much less systematic in its presentation and more difficult to singularly interpret than Fichte's work, can be examined in precisely the way in which Fichte understood it.

2 Fichte's notion of drive

Fichte's early notion of drive should be understood in the context of his developing a philosophical account that directly responds to the ending of Karl Leonhard Reinhold's *Versuch einer neuen Theorie des menschlichen Vorstellungsvermögens* (1789). When comparing this book with Fichte's *Versuch einer Kritik*

³ See Herder's "voll unwiderstehlichen triebes reden sie" versus his "eigenen platz soll er finden [der baum], damit er durch eignen trieb wurzelaus in die höhe steige": *Wörterbuch*, 446 and 442, respectively.

aller Offenbarung (1792) it is striking that, in both books, the discussion of drive leads to pure will and freedom (cf. Reinhold 1789, 567–572; GA I,I, 31–38). In both cases, drive means a drive to representation, which splits into a matter drive and a form drive.⁴ Although drive leads to the will for both authors, they refer to it purely as an epistemic process that facilitates reflection. Fichte essentially repeats this Reinholdian use of the volitional drive as leading towards freedom in the *Grundlage der gesamten Wissenschaftslehre* (1794/95).

In Fichte's *Einige Vorlesungen über die Bestimmung des Gelehrten* (1794) we find a noticeably different way of discussing drive, here not limited to representation, but as a “Grundtrieb,” “höchste Trieb” and “gesellschaftliche Trieb” (GA I,III, 35, 38, 39). In his *Versuch*, Fichte had connected the representational drive to will and sociability, but the notions of a foundational and highest drive are, at this time, wholly new in his thought. In order to explain the importance and origin of these terms, we will consider his extended account of drive offered in the lecture transcripts of *Vorlesung über die Moral* (1796). Some of these conceptual distinctions can be also be found in the third part of the *Grundlage*, but not in reference to terms like “Grundtrieb” and not with the same complexity of argumentation.

Throughout these lectures, Fichte examines the transition from unconscious processes to consciousness and thereby allows for a rational reconstruction of the way in which these unconscious processes affect, or in some cases even constitute, our conscious reflection. Fichte's central argument is that there is a volitional component⁵ to all activities of the mind, one which furthermore precedes and makes possible any and all determinations of the will. Incorporating the earlier epistemic use of drive from the *Versuch*, he argues that it is this very volition that produces our cognitions and representations and the resulting “feeling of reality” (GA IV,I, 19) that accompanies them, further demonstrating that cognition and action are essentially and inextricably related (GA IV,I, 53).

⁴ One can also locate Friedrich Schiller's use of drive in *Briefe über die Ästhetische Erziehung des Menschen* (1795), which was later expanded on in *Kleine prosaische Schriften*, Band 3 (1801), in the category of uses that correspond directly to Reinhold. In relation to Reinhold, Schiller adds the “Spieltrieb” by way of mediation, but does not transform the semantic context of the term.

⁵ Fichte sometimes uses “Trieb” and “Tendenz” interchangeably. Although his use does not always seem to be entirely consistent, on many occasions he refers to “absolute tendency” as the initial impetus of cognition and action, and refers to “drive” for the cases in which this process is accessible to reflective consciousness. Cf. GA IV,I, 39. All translations of Fichte's texts are my own.

Fichte rejects any causal model that might be included in a notion of drive. Whereas botanical and zoological senses are inextricably tied to a causal account, Fichte is adamant that his notion of drive must function as the absolute limit of any explanation. That is to say, it explains everything in our mental world, but can itself not be explained. Fichte argues that drive should be thought of as a law, rather than as part of a mechanical or causal system (GA IV,I, 25).⁶ The notion of drive that Fichte puts forward is thus essentially constitutive for human cognition and action, not in the last place because we cannot provide an explanation of drive beyond its effective scope within human activity. As a result, drive functions as the primary operation, one that consequently prevents us from understanding the world around us as a predetermined universe. After all, the inconceivable and inscrutable drive is the source of all cognition, offering a merely incomplete causal model which draws attention to the human being as a free agent.

Rather than attempting to reconcile freedom and determinism, Fichte, throughout his career, reframes the knowable world as the result of human activity. In this respect, we must understand his notion of drive as a component of the utmost importance, because it links the relatively impersonal idealistic world-view to our free agency (GA IV,I, 49). What facilitates this linking however, is our ability to *reflect* on our drive.

Fichte cautions that this means that we have to limit ourselves to goals that are *attainable*, for instance, in the sense that our goals relate to experience. If we do not set our goals within experience, which was the result of cognition volitionally produced by our drive, we are still, in Fichte's words, "slaves to the laws of nature" (GA IV,I, 38). Why would this be the case? Because, if we do not, in our actions, utilize the cognitions that drive offers us, these cognitions will consequently only impede our actions, because in that case these cognitions are illegitimately considered to be the *necessary* limits of our actions. If we do not set attainable goals, we do not reflect on nature and subsequently lose our ability to extricate ourselves from it. The end result would be a kind of reality conservatism, in which we cannot think the world in a different way than it is, because we do not understand it and thus cannot conceive of the rational means required to effectively interact with it. The human being must thus, if he is not to deny his drive as the source of his cognitions, reflect on his cognitions and thereby create a rational way to have agency in relation to these cognitions.

This leads us to Fichte's articulation of the difference between the empirical and the transcendental perspective. The empirical perspective understands drive

6 Also see Fichte's: no ground can be "given" for "absolute drive" (GA IV,I, 35).

as part of a mechanism in nature [Naturmechanismus], in which the human being has a volition because it has been caused to have this volition by a system of nature (GA IV,I, 41). The transcendental perspective, which is Fichte's perspective, shows that the empirical perspective is erroneous because drive is self-determinative, rather than causally determined (GA IV,I, 42). The empirical perspective on drive leads one to conclude that drive is part of the subject in as far as it is a part of nature (GA IV,I, 48). The transcendental perspective recognizes that the human being is free in the sense that he is, through his drive, disconnected from natural causality (GA IV,I, 49).

Whereas both types of consciousness *have* drive, strictly speaking, the degree of consciousness that we have *of* that drive opens the door towards the transcendental perspective. Reflection, in Fichte's words, "loosens man from his animal instinct" (GA IV,I, 51). Fichte uses the word "instinct" to refer to those processes that are of drive, but which remain unconscious. For instance, drive is the common denominator of volition in all life forms: plants, animals and humans. Fichte classes the processes that function without our being conscious of it as instinctual. He therefore proclaims: "Man has animal instinct, but he also does not have it." (GA IV,I, 51) The reason for this ambiguity can be found in the fact that we have the potential to become conscious of all of our animal desires, thereby, in a sense, *negating* our animal nature, because these desires inevitably become moments of genuine *choice*, in which we choose to determine our will according to them or not. For Fichte, drive becomes a powerful conceptual tool that allows him to discuss the potential scope of human action, and the agency that we can reclaim from animal instinct by virtue of reflection.

One might wonder how reflection achieves this remarkable feat. Fichte rejects Moses Mendelssohn's and Johann Friedrich Wilhelm Jerusalem's common claim that the human being simply aims to improve itself. Fichte claims that reflection introduces a general model of desire into our rational deliberation, which he calls *yearning* (*sehnen*). It is this general model which aids the human being in becoming aware of desires (GA IV,I, 49–50). Presumably it is this model that allows one to register animal desires as desires which are amendable and can be made subject to our will. This general model of desire also involves a general kind of pleasure, which Fichte calls lust [Lust], and which presumably motivates our reflection about practical matters (GA IV,I, 52). One can draw parallels to Kant's disinterested interest, since Fichte offers a type of contemplative motivation that is unrelated to any specific object. This model of yearning and lust allows the human being to formulate and implement specific desires according to our will.

In addition to animal instinct, Fichte also briefly outlines plant instinct, which is, somewhat surprisingly, also part of the human being. Plant instinct

is that set of processes which we *cannot* submit to our will by virtue of reflection, and which we therefore cannot effectively impact through choice (GA IV,I, 51). The sole example that Fichte provides in this instance is the process of our digestive system. Reflecting on these processes does not directly enlarge our scope of action, because we can scarcely *will* to modify or cease our digestion. This is one of the hard limits of human action within Fichte's account (GA IV,I, 18).

3 Jacobi's notion of drive

Although there are no overt references to Jacobi in Fichte's lectures, the commentator on the transcript, who is likely Otto Heinrich von Mirbach (GA IV,I, 3), notes the obvious conceptual and argumentative similarities to Jacobi's position (GA IV,I, 51, 116). I argue that the commentator is correct, and that our understanding of the comparison will be better served by a careful reconstruction of the notion of drive that Jacobi put forward in the texts that were available to Fichte. Although Fichte developed many of the arguments that made the new philosophical notion of drive attractive, the core idea was clearly developed by Jacobi.

I argue that three texts are relevant here. The first text is the 1789 second edition of *Über die Lehre des Spinoza in Briefen an den Herrn Moses Mendelssohn*, in particular the new introductory text and appendixes. These additions are meant to dispel the idea that Jacobi argued for a religious enthusiasm in the first edition, and develop the epistemological ideas in *David Hume über den Glauben, oder Idealismus und Realismus* (1787) in relation to ethics and anthropology. This second edition was highly praised by both Kant and Reinhold, who previously had opposed the first edition. Many of the ideas developed in these additions are developed in two other texts, Jacobi's two novels *Eduard Allwills Briefsammlung* and *Woldemar*. During the early 1790's Jacobi rewrote these novels to reflect the development of his philosophical views, publishing new versions of *Eduard Allwills Briefsammlung* and *Woldemar* in 1792 and 1794, respectively.⁷

In the 1789 text, we find that Jacobi is concerned with the notion of a drive that can connect instinct and human choice, and consequently functions as a hinge between understanding the human being as part of a natural mechanism,

⁷ It is only during the late 1780s that Jacobi developed his new notion of drive. In *Kunstgarten* (1781), a prose piece that would later be incorporated in *Woldemar*, Jacobi uses the literary and zoological sense, and tries to relate these through the volitional sense, but does not yet have a philosophical vocabulary that is sophisticated enough to bring these senses together systematically (JWA 7,1, 124–127, 172–175). All translations of Jacobi's texts are my own.

and, on the other hand, as a free agent (hence the titles of the two sections: “Man is not free” and “Man is free” in the introductory essay).⁸ Jacobi remarks that all the various desires of living beings are founded on an “original natural drive,” whose function it is to “preserve and enlarge” the natural being’s “ability to exist” [da zu sein] (JWA 1,1, 159). This “a priori of desire” is also the original drive of the rational being. Like Fichte after him, Jacobi combines the zoological and the volitional drive, establishing drive as the condition of the possibility of human action.

Jacobi also recognizes that the rational being introduces something new into this natural framework, because it has a “higher degree of consciousness,” leading to the fact that it is conscious of its identity, or rather is self-conscious, as we would now put it (JWA 1,1, 159). Subsequently, the modification of the natural drive in the rational being is that, due to the introduction of reflection, in this case self-consciousness, it seeks to enlarge its degree of *personhood* rather than merely its existence (JWA 1,1, 159). The “rational desire” (JWA 1,1, 159) that follows directly from this drive is what Jacobi calls the *will*. We can see the same conditional structure here that we also saw in Fichte’s arguments. The advantage of framing the will as a rational desire, thereby classing it as something derived from drive, is that one no longer needs to explain the origin of a pure will. For Jacobi, the rational desire is only different from other desires because of the introduction of rationality, which introduces the possibility of self-consciousness.

It should be noted that Jacobi distinguishes rational desires, which are principled determinations of the will, from non-rational desires. This means that the natural drive can *interrupt* the personality that is enlarged through the rational drive (JWA 1,1, 160). Like Fichte, Jacobi treats these interruptions as regressions: the rational being surrenders a part of itself because it surrenders a potential object of reflection. On the other hand, the rational being must also consider the non-rational desires as its *own* desires, in that it is *accountable* for them (JWA 1,1, 160). Here, again, we find a similarity with Fichte, who objects to the idea that following a natural drive might absolve us of the responsibility for our agency, in the sense that we might be able to blame our misdeeds on a natural causality working through us. Yearning, as a general concept for desire, introduces the possibility of guilt for actions that follow from animal desires (GA IV,I, 105).

⁸ Jacobi added this text to his infamous open letter to Fichte, likely to show Fichte that Jacobi had caught on to the fact that Fichte was, in many ways, greatly indebted to him.

As Jacobi put it: “The entire system of practical reason is based on this experience, *in as far as it is built on one foundational drive*.” (JWA 1,1, 160)⁹ Very likely, it was this phrase that struck a chord with Fichte. It presents an account of free agency that does not neglect non-rational desires, by way of introducing a “foundational drive” [Grundtriebe]¹⁰ of which both natural and rational drive are a result. It also presents a mandate for the rational being to progressively reflect on his natural desires and relate them to the foundational drive so as not to vilify the impulse behind them. Jacobi’s position thus contains Fichte’s position in outline: a rational being has a sense of agency thrust upon him by drive and gains the capability to reflect on this entire process. This characterization prefigures the line that Fichte would draw between animals and humans: the human being is distinct from the animal in as far as he reflects on his position, allowing him to become completely distinct from animal desire and natural causality.¹¹ Like Fichte, Jacobi argues that the foundational drive is conditionally implicated in rationality (JWA 1,1, 162).

In line with the incorporation of non-rational desires within a reflective framework, Jacobi argues that the rational being must “conquer” [unterwerfen] that which injures this personal identity as it is constituted by the reflective

9 Jacobi grounds our sense of justice on this structure, remarking that this inner sense of justice mechanically emerges in any human being with a consciousness of his identity (JWA 1,1, 161). He further argues that this sense of justice applies, *in abstracto*, to all human beings in exactly the same way as it would to oneself, because the rational being, as a rational being, has no grounds to distinguish between the two: “*I and man are one. He and man are one. Therefore, I and he are one.*” (JWA 1,1, 161)

10 Since Jacobi offers the earliest use of the “Trieb” term we have covered, a short remark on its history is in order. “Grundtrieb” has a slightly different meaning than “Trieb” in general, because it was used in publications during the 1760s and 1770s. The most prominent example is Johann Georg Heinrich Feder (Feder 1769, *Grundriß der philosophischen Wissenschaften*). Feder uses the term to refer to a primary moral passion. Feder rejects the account that the “drive to cognition” [Erkenntnißtrieb] is “the one true and absolute foundational drive” (Feder 1779, 228). One could say that these authors attempt to formulate a comprehensive volition of human nature, but lacked the conceptual tools to produce a comprehensive account. Johann Georg Sulzer uses “Trieb” in a similar way (Sulzer 1773, 37). Incidentally, Sulzer was one of Jacobi’s inspirations. In a way, Jacobi and Fichte take up the open challenge to present an integrative notion of drive that covers epistemic, naturalist and practical dimensions. Although Fichte refers to Johann Friedrich Wilhelm Jerusalem in these lectures, who also uses “Grundtrieb,” Jerusalem uses it in an extremely limited sense, as self-love and a drive to perfection (GA IV,I, 97, 119). Jacobi is a much more likely source for Fichte, because he shares Fichte’s aim of an integrative and dynamic notion of drive.

11 As we’ve seen, Fichte adds plant instinct and, subsequently, the potential for a class of actions which is not accessible to reflection and therefore adds no scope or content to our free agency.

act (JWA 1,1, 161). He thereby locates the feat of overcoming base desires within our capacity to reflect on them. This line of reasoning introduces a volitional component to reflection itself, which paved the way for Fichte's later incorporation of the entire epistemic process within his account of drive. Put in the context of the entire book, Jacobi's account of drive is meant as a solution to the abstract reasoning that he criticized in the first edition and to maintain an applied account of rationality (JWA 1,1, 166).¹² We can understand Fichte as taking up a very similar project, which finds its clearest expression in his account of drive.

It is highly likely that this text was of decisive importance for the development of Fichte's notion of drive, not only due to the conceptual similarities with Fichte's later project, but also due to the conceptual *complexity* that belies Jacobi's 1789 account of drive.¹³ While many of these conceptual distinctions also occur in 1794's *Woldemar*, the absence of clear argumentation makes it an unlikely candidate for transmission.¹⁴ There Jacobi further develops his views on virtue in relation to drive, but this particular ethical theory does not wholly resemble Fichte's views on the matter, although the notion of drive is very similar. Likewise, the amount of conceptual material in 1792's *Allwill* is limited, but it might have been an additional source of inspiration for Fichte, since Jacobi refers to a "pure drive" in that text, as the productive power of the mind which fills the senses with determinations (JWA 6,1, 224). This further adds to the idea that our sensible cognition is generated wholly for the benefit of our drive. When read together with the 1789 account, this text offers the full sense of Fichte's notion of drive, wherein epistemology operates in tandem with rational action. Whereas we can likely credit Jacobi with developing the concept towards its integrative, more philosophically inclined usage from the late 18th century onwards, Fichte was instrumental in developing the systematic presentation associated with drive.

4 Conclusion

A full account of the development of this integrative and comprehensive notion of drive, which offers the volition for every human activity, should also involve a comparison of the 1796 lectures on morality with the *Grundlage der gesamten*

¹² "Human reason, isolated from humans themselves and from all drive, is a mere thing of thought [Gedankending], that cannot operate [agieren] nor react, neither think nor act." (GA 1,1, 166)

¹³ One passage, for instance, reads like a commentary on Jacobi's text (GA IV,I, 47).

¹⁴ For instance, it uses natural drive and fundamental drive (GA 7,I, 225, 249).

Wissenschaftslehre, the *System der Sittenlehre nach den Principien der Wissenschaftslehre* and the *Wissenschaftslehre nova methodo*.¹⁵ That task must be undertaken in a different publication. For instance, there are clear indications that the lectures were the basis of the *Sittenlehre*, while the lectures offer a lot more clarity on the topic of drive. Only an extensive comparison of the texts can discover the reasons behind Fichte's choice to restrict his extended account of drive to his lectures.

In response to Jacobi's and Fichte's drive, realist authors like Johann Neeb (Neeb 1817, 248; Neeb 1796, 382) and Wilhelm Traugott Krug (Krug 1803, 204–5) use a similar notion, further contributing to the rise in the usage of "Trieb." Much can be gained by looking at the development of this notion from Fichte onwards, towards the early romantics, Schopenhauer and Freud – in this sense this chapter offers only a first glance at the emergence of the notion of drive that would be so influential throughout the 19th century.¹⁶

Jacobi and Fichte have poised drive as a hinge between naturalism and ethics in a way that is still relevant for contemporary discussions. One need only consider discussions on the role of the introduction of consciousness into a determinist account in order to see what can be gained from a notion that does not decisively favor either of these positions, while retaining the pragmatic uses of both approaches. In both Jacobi's and Fichte's account, drive integrates various aspects of the individual and facilitates a unified account of experience, which in turn facilitates rational deliberation.

A volitional sense of drive structures the new notion and allows for the increased possibilities of application. In Fichte's account this includes the botanical and the zoological sense (as plant and animal instincts), but also the literary sense in as far as self-consciousness is involved, in the sense that yearning, as an abstract model of desire, is exactly this striving towards an undetermined object that is so central to the human experience. Jacobi and Fichte are mostly in agreement, although Jacobi does not specifically make a distinction between the botanical and zoological sense, preferring to refer to both as instinct.

¹⁵ See Manja Kisner's chapter *Drive and Will in Fichte's System of Ethics* in this volume for an examination of the *Sittenlehre* from the perspective of drive.

¹⁶ Schopenhauer capitalizes greatly on the epistemological drive as *unconscious*, whereas Freud developed a theory that was initially more in line with the zoological sense towards a model that aims to explain social developments as *unconscious*, further developing the connection between psychology, epistemology and ethics that Jacobi and Fichte first developed in relation to drive.

One might be tempted to consider the role that drive has in making possible reflection¹⁷ and, subsequently, freedom as an emergent quality, similar to Nicolai Hartmann's "Novum" (Hartmann 1940). It might also be worthwhile to compare Jacobi's and Fichte's drive with Georg Wilhelm Friedrich Hegel's notion of spirit in this respect.

Finally, with the inclusion of plant instinct, Fichte opens the door to a class of unconscious processes that we might be able to reflect on abstractly, but which cannot directly enlarge our scope of action. There are then two types of unconscious processes involved in his account of drive: the not yet conscious of animal instinct and the non-conscious of plant instinct, in effect also extending the scope of non-conscious processes that can still be considered part of the human being's activities.

Bibliography

- Feder, Johann Georg Heinrich (1769): *Grundriß der philosophischen Wissenschaften*. Coburg: Verlegts Johann Carl Findeisen.
- Fichte, Johann Gottlieb (1962–2012): *Gesamtausgabe*. Stuttgart-Bad Cannstatt: Frommann-Holzboog. (cited as GA)
- Forster, Georg (1783): *Bemerkungen über Gegenstände der physischen Erdbeschreibung, Naturgeschichte und sittlichen Philosophie auf seiner Reise um die Welt*. Wien: Trattner.
- Grimm, Jacob (1952): *Deutsches Wörterbuch von Jacob Grimm und Wilhelm Grimm*, Band 22. München: Deutscher Taschenbuch Verlag.
- Hartmann, Nicolai (1940): *Der Aufbau der realen Welt. Grundriss der allgemeinen Kategorienlehre*. Berlin: Unknown.
- Jacobi, Friedrich Heinrich (1787): *David Hume über den Glauben, oder Idealismus und Realismus Ein Gespräch*. Breslau: Gottl. Löwe.
- Jacobi, Friedrich Heinrich (1789): *Über die Lehre des Spinoza in Briefen an den Herrn Moses Mendelssohn*. Breslau: Gottl. Löwe.
- Jacobi, Friedrich Heinrich (1812): *Werke*, Band 1. Leipzig: Fleischer.
- Jacobi, Friedrich Heinrich (1998–2016): *Jacobi Werke Gesamtausgabe*. Meiner: Hamburg. (cited as JWA)

¹⁷ This, of course, does not really explain how reflection emerges in a process that is initially a natural causality. One could argue that it is exactly this problem that makes Fichte propose reflection as an emergent quality, because he incorporates both the causal process and the position of freedom, making it clear that there is a connection between the two for reflective consciousness, if one reconstructs the causal account, but that this connection cannot be understood without fully depending on the "novum" of reflection. Fichte tries to explain reflection by referring to the fact that we are subject to a "radical evil," a "delay" in our response to input (GA IV,I, 98). This does not so much explain the emergence of reflection, but rather how it functions.

- Krug, Wilhelm Traugott (1803): *Fundamentalphilosophie*. Züllichau: Darnmannsche Buchhandlung.
- Neeb, Johann (1796): *System der kritischen Philosophie auf den Satz des Bewustseyns gegründet*, Band 2. Bonn: Andreäischen Buchhandlung.
- Neeb, Johann (1817): *Vermischte Schriften*, Band 2. Frankfurt a.M.: Hermann.
- Reinhold, Karl Leonhard (1789): *Versuch einer neuen Theorie des menschlichen Vorstellungsvermögens*. Prag, Jena: Widtmann und Mauke
- Schiller, Friedrich (1801): *Kleine prosaische Schriften*, Band 3. Leipzig: Siegfried Lebrecht Crusius.
- Sulzer, Johann Georg (1773): *Vermischte philosophische Schriften*, Band 1. Leipzig: Weidmanns Erben und Reich.

Manja Kisner

Drive and Will in Fichte's *System of Ethics*

Abstract: In this paper, I argue that the concept of will in Fichte's *System of Ethics* (1798) is closely connected to the concept of drive (*Trieb*). To show this, I first explicate Fichte's notion of drive and after that the relation between the drive and the will. In Fichte's system, the concept of drive functions as an intermediary link between the domain of nature and freedom and as such is central for explaining Fichte's theory of the will. The link between these two concepts enables Fichte to describe the will in dynamic terms and to point out the activity of the will.

1 Introduction

In recent years, Fichte's philosophy has experienced increased interest in the Anglophone world. The focus of this revival is Fichte's moral philosophy.¹ In 2016, Allen Wood published a monograph on *Fichte's Ethical Thought*, in which he characterizes Fichte as “the most influential single figure on the entire tradition of continental European philosophy in the last two centuries,” but also the most neglected (Wood 2016, ix). Despite this latter verdict, Fichte's philosophy, and in particular his ethical theory, continues to gain attention. In 2018, another monograph – Michelle Kosch's *Fichte's Ethics* – and new papers dealing with Fichte's ethical thought appeared.² In this paper, I will follow this trend by analyzing and discussing a particular theme from Fichte's moral philosophy, namely the relation between the concept of will and the concept of drive in his *System of Ethics*.

The concept of will plays a crucial role in Fichte's *System of Ethics*.³ Fichte follows Kant in this regard, who emphasized the significance of this concept

¹ Fichte's older works in English include a number of papers from Daniel Breazeale. For the discussion of Fichte's practical philosophy see for instance Breazeale 1996; Breazeale 2006; Breazeale 2012. Moreover, in 1998 Günter Zöller published a monograph in English on the original duplicity of intelligence and will. See Zöller 1998. In German language the reception of Fichte's work was steadier than in English literature, although Fichte is not studied and discussed as much as Kant or Hegel.

² In the last years, especially Owen Ware has published extensively on Fichte's ethics (see Ware 2010; Ware 2015; Ware 2018).

³ References to Fichte's texts are made by citing the volume and page number of Fichte's *Gesamtausgabe* (GA). All translations are from Fichte 2005 (translated and edited by Daniel Breazeale and Günter Zöller).

to his moral philosophy as well. At the time Fichte was writing his *System*, he was well acquainted with Kant's early works on moral philosophy, including the *Groundwork of the Metaphysics of Morals* (1785) and the second *Critique* (1788). He also read Kant's *Religion within the Boundaries of Mere Reason* (1793), which is relevant for understanding the distinction between *Wille* and *Willkür*. Only Kant's later *Metaphysics of Morals*, published in 1798, the same year as the *System*, could not have influenced Fichte's book, especially because we know that already from 1796 onwards Fichte was lecturing on these topics.⁴ It is beyond doubt that Kant's works on moral philosophy – with the exception of his *Metaphysics of Morals* – were of great importance to Fichte, who saw himself as a follower of Kant and his *System* as a work that expanded on Kant's practical philosophy.

Moreover, with regard to the concept of drive we can also observe a Kantian influence on Fichte.⁵ It is not only Kant's practical philosophy that deeply inspired Fichte, but also the third *Critique* with its discussion of the formative drive (*Bildungstrieb*) and the concept of organism.⁶ In contrast to Fichte, however, Kant does not use the concept of drive as a part of his ethical theory.⁷ The in-

4 Fichte's *System* belongs to the second phase of Fichte's *Wissenschaftslehre*, namely to his *Wissenschaftslehre nova methodo*, on which he was lecturing between 1796 and 1799. For the preceding discussion of the concept of drive in Fichte's lecture on morality (*Collegium über die Moral*, 1796) see Giesbers' chapter *Drive as a Constitutive Element of Practical Action in Jacobi and Fichte* in this volume.

5 Of course, Kant is not the only influential figure for Fichte. With regard to the connection between the will and the drive, Reinhold's distinction between selfish (*eigennütziger Trieb*) and unselfish drive (*uneigennütziger Trieb*) established in his *Letters on the Kantian Philosophy, Volume II* is important. Moreover, it is plausible to assume that Jacobi's concept of drive was relevant to Fichte as well. See Giesbers' chapter in this volume.

6 Actually, it was Kant's third *Critique* that Fichte read first and only after that proceeded to Kant's practical philosophy and the second *Critique*.

7 In Kant's practical philosophy, in contrast to Fichte's *System*, the word drive has the same or at least a very similar meaning as the word inclination (*Neigung*); he uses the term inclination first to emphasize the gap between morality and nature. For Kant there are moral duties on the one hand and inclinations that do not belong to the moral domain on the other. Cf. MS, AA 06: 215–216: "Only the natural drives (*die natürlichen Triebe*) for food, sex, rest, and movement, and (as our natural predispositions develop) for honor, for enlarging our cognition and so forth, can tell each of us, and each only in his particular way, in what he will find those joys; and, in the same way, only experience can teach him the means by which to seek them. [...] But it is different with the teachings of morality. They command for everyone, without taking account of his inclinations (*Neigungen*), merely because and insofar as he is free and has practical reason." Drives or inclinations belong to nature and sensibility; morality, by contrast, has to be free of natural drives. Therefore, Kant's notion of the drive understood as sensible inclination only

tegration of drives and feelings into the system of ethics is specific to Fichte's approach, which starts with a Kantian background but develops into a new direction that we cannot find in Kant.⁸ Fichte's *System* offers a synthesis of nature and freedom, one to which Kant was aspiring in his third *Critique*, but which Fichte tries to establish in the domain of his ethical theory. In this step Fichte goes in another direction than Kant, who sharply distinguishes between freedom as a domain of moral philosophy and another domain of nature.

In this paper, I discuss the relation of the will to the concept of drive in Fichte's *System*. I argue that the link between these two concepts enables Fichte to interpret the will as a dynamic concept and to see it as a manifestation of the activity of the I. This dynamic aspect of the will is something original to Fichte's understanding and is missing in Kant's conception. In Section 2, I introduce Fichte's concept of will as an act of willing. In Section 3, I give the preliminaries to his concept of drive. Section 4 then examines in more detail Fichte's system of drives and focuses on the intermediary role that these concepts play in connecting freedom and nature. In Section 5, I return to Fichte's concept of will and point to the connection between the system of drives and the concept of will. In this way, I demonstrate the way in which examination of the concept of drive helps us to explain Fichte's will as a dynamic concept.

2 The concept of will and the tendency to self-activity

Throughout the *System* Fichte highlights the key role of the concept of will for his moral philosophy, but he nonetheless does not explain and define this concept unambiguously. Fichte uses different characterizations of the will and he also distinguishes – by following Kant in his *Religion* as well as Reinhold in his *Lectures on the Kantian Philosophy Volume II* – between *Wille* and *Willkür*. In this section, I will focus on the meaning of Fichte's concept of *willing* and specify in which regard this concept differs from Kant's understanding of the will (for further discussion on the distinction between *Wille* and *Willkür* see Section 5).

plays a negative role in Kant's moral philosophy: Kant refers to the drive to point out what cannot be a part of his moral theory.

⁸ Fichte employs the concept of drive already prior to the *System*: already in the *Foundations of the Entire Science of Knowledge* (1794/95) as well as in his lectures on *Wissenschaftslehre nova methodo* Fichte discusses at length the concept of drive and relates it with the concept of striving.

Even though we must insist on the crucial differences between Kant's and Fichte's approach to ethics, it is nevertheless true that Fichte adopts many of Kant's central concepts and develops his system against the background of Kant's practical philosophy. To start with one relevant example, Fichte clearly follows Kant in the introduction to the *System* when he relates the concept of will to the concept of an end (*Zweckbegriff*):

Something objective follows from a concept. How is this possible? What can it mean? It can mean only that the concept itself appears to me as something objective. But the concept of an end, viewed objectively, is called an act of willing [*ein Wollen*], and the representation of a will [*eines Willens*] is nothing other than this necessary aspect of the concept of an end, which is posited solely in order to make possible the consciousness of our activity. What is spiritual in me, intuited immediately as the principle of an efficacious action, becomes for me a will. (GA I/5, 28)

Here, Fichte identifies the concept of an end with an act of willing (*ein Wollen*) by claiming that when we try to represent the concept of an end objectively, we must conceive it as an act of willing. Willing is hence a term that denotes an objective manifestation of something that is originally not objective. Thus, the distinction between an objective and a subjective viewpoint is of immense importance for Fichte. The I, viewed from a subjective perspective, produces a concept of an end and therefore this concept is a product of an act of representing.⁹ The concept of willing, by contrast, is an objective manifestation of the concept of an end.¹⁰

⁹ Cf. GA I/5, 90: "Willing is thought of as preceded by a freely active comprehending of an end, that is, by an absolute producing of the end by means of a concept. In producing the concept of an end, the state of the I is purely ideal and subjective. Something is represented, and it is represented by means of absolute self-activity, since the concept of an end is only a product of an act of representing; something is represented in relation to a future act of willing, for otherwise the concept [that is represented] would not be a concept of an end; but what is represented is only represented and is by no means willed."

¹⁰ These two perspectives, the subjective and the objective, are also relevant to other relations between concepts, and not only to those between the concept of an end and willing. For instance, Fichte also uses the distinction between the subjective and objective perspective when he relates the body and the will. In this case, however, willing represents the subjective aspect and the body is viewed as an objective articulation of the will. Cf. GA I/5, 28f: "But now the will is supposed to exercise causality, and indeed, an immediate causality upon my body; and the body as an instrument, that is, the articulated body (*die Artikulation*), extends only as far as this immediate causality of the will extends. [...] The will is therefore also different from the body, and it appears as not being the same as the body. This distinction, however, is nothing more than yet another separation of what is subjective from what is objective, or more specifi-

Despite Fichte's different terminology, his definition of the concept of an end nonetheless resembles Kant's definition. Kant takes the concept of an end to denote a special kind of causality, namely the causality of concepts, which is constitutive for his practical philosophy and represents an alternative to his conception of natural causality in his theoretical philosophy. Moreover, he relates the concept of an end with the will, as the following examples from the *Groundwork* and the third *Critique* illustrate:

Now, what serves the will as the objective ground of its self-determination is an end (*der Zweck*), and this, if it is given by reason alone, must hold equally for all rational beings. (GMS, AA 04: 427)

The faculty of desire, insofar as it is determinable only through concepts, i.e., to act in accordance with the representation of an end, would be the will. (KU, AA 5:220)

According to Kant, we can speak about the causality of concepts when the concept determines the will and serves as its objective ground. The causality of concepts is hence a special causality that does not determine the objects of nature, but only our will. This understanding corresponds very well with Fichte's claim from the quotation above that "the concept of an end, viewed objectively, is called an act of willing" (GA I/5, 28). In both cases, the will is conceived of as an objective manifestation of the efficacy of concepts.¹¹

But there is also a crucial difference between Kant and Fichte that concerns Fichte's dynamic character of the will. For Fichte, as I will argue, the essence of the will lies in its activity (*Tätigkeit*). Although important to Fichte, this dynamic aspect is not explicitly emphasized in Kant's conception of the will. In the first part of the *System*, Fichte dedicates the deduction of the principle of morality to proving that the essence of the I is pure activity and that consequently the concept of will must also be conceived of as an objective manifestation of this activ-

cally, it is a particular aspect of the original separation. In this relationship the will is what is subjective and the body is what is objective."

11 Accordingly, Frederick Neuhouser interprets Fichte's will as being determined by a purposive concept and hence being a result of the causality of a concept. Cf. Neuhouser 1990, 126: "The kind of concept Fichte has in mind here is what he calls a *Zweckbegriff*, or 'purposive concept.' Thus, to say that the 'absoluteness' of the practical subject 'comes under the jurisdiction of concepts' must be to claim that the will's transition from its original indeterminacy to a determined state (i.e., to willing a particular something) takes place in accord with the subject's conception of the purpose it wants to pursue through its action. [...] Furthermore, it is only when such a purposive concept is involved in the will's determination of itself that its action is genuinely free, as opposed to merely uncaused, or undetermined."

ity. Fichte describes this activity of the I in two ways, first with regard to thinking and then with regard to willing.

The I, viewed subjectively, is a pure activity or self-activity (*Selbsttätigkeit*). In this first sense, the I is conceived of as consciousness and hence corresponds to the faculty of thinking. In the second sense, this is from an objective perspective, this activity appears as willing.¹² Thinking, according to this view, is “for itself, by no means an object of any particular new consciousness but is simply consciousness itself” (GA I/5, 39). Fichte regards willing, by contrast, as “an originally objective manifestation of the substance in question here” (GA I/5, 39) and thus calls the I a substance whose manifestation we experience as willing. As such, willing “always remains only objective and is never itself an act of thinking but is always only the manifestation of self-activity (*Selbsttätigkeit*) insofar as the latter is thought” (GA I/5, 39).

This second viewpoint is important if we want to reflect on the I as a pure activity. Consequently, Fichte starts his deduction with the claim that we are acquainted with ourselves as willing beings: “I find myself as myself only as willing” (GA I/5, 37). The result of this understanding is that it is only through willing that we become conscious of ourselves: “In short, the sole manifestation [of the substantial I] that I originally ascribe to myself is willing. Only under the condition that I become conscious of willing do I become conscious of myself” (GA I/5, 39).

As we see above, in this context Fichte calls the I a substance. However, the subsequent explanation shows that Fichte’s I cannot be understood as a substance in the traditional use of the term, as something persisting and unchangeable.¹³ On the contrary, the dynamic aspect of the will makes it impossible to interpret the I as a substance. For Fichte, the concept of will is an objective expression of the internal activity of the I and as such means something different than Kant’s concept of will.¹⁴

12 Therefore, Fichte’s original, absolute I splits in two – into the subjective and the objective I, as two different ways of looking at the I. The objective I, conceived of as willing, and the subjective I, conceived of as thinking, represent two different sides of one and the same absolute I. Cf. Metz 2006, 26 f.

13 It is for this reason that Frederick Beiser understands Fichte’s concept of the I and his conception of self-positing in opposition to not only the previous philosophical traditions, but also to Kant. Cf. Beiser 2002, 281.

14 Allen Wood points to this difference between Kant’s and Fichte’s conception of the will. Cf. Wood 2016, 67: “For Kant, the free will is a substance, having causal powers like other substances; according to the mechanism of nature, the causal powers of substances must be grounded in other, prior causal powers, which necessitate its causal action. The animal power of choice (tierische Willkür, arbitrium brutum) is a will of this kind. Free will (arbitrium liberum) must

Fichte then further describes this activity as a striving or a tendency of the I: “The essential character of the I, through which it distinguishes itself from everything outside of it, consists in a tendency to self-activity for self-activity’s sake” (GA I/5, 45). This tendency is a “tendency to determine itself absolutely, without any external impetus” (GA I/5, 45). And in this way, with his notion of self-determination, Fichte comes close to Kant’s autonomy, although in this case Fichte’s notion also differs from Kant’s in regard to its dynamic aspect: What Fichte points out is only the tendency to self-activity and self-determination and in this tendency Fichte’s I is not a static concept but characterized through its endless activity. As a result, Fichte defines willing as “a real determining of oneself through oneself” (GA I/5, 40).

To summarize, we can maintain that Fichte’s concept of will brings to expression this inner tendency of the I, indicating the dynamic nature of the I. However, in the first part of the *System* Fichte does not limit his explanation of the dynamicity of the I only to his discussion of the will as an act of willing; rather, he also points to another term that can stress the dynamicity of the I even more clearly. For this task, Fichte introduces the concept of drive to the analysis, to which I will now turn.

3 Preliminaries on the concept of drive

Even though Fichte recognizes – as I have shown in the previous section – the tendency to self-activity for self-activity’s sake as an essence of the I, he also must admit that we cannot conceive of this tendency directly. Instead, the tendency of the I is something we come to only after abstracting from everything that is foreign to the I. Therefore, he proposes another way to explain this tendency – namely with the help of the concept of drive. For this reason, Fichte refers to the concept of drive in the first part of the *System* and then in even more detail in the second part, in which he develops a whole “system of drives and

be a substance which acts ‘from itself,’ and is an exception to the natural mechanism. Fichte, by contrast, holds that the will or willing I is only an acting. Acts of will are by their concept free; they are not part of the causal mechanism of nature at all, though motions of the human body, which are the objective side of these acts, do cause certain effects in the material world. For Fichte, it is a conceptual truth that all willing is free: ‘an unfree will is an absurdity’ (SL 4:159). Fichte also differs from Kant in that he does not regard the state that is begun absolutely as begun solely ‘from itself’ in the sense that it is unconnected with something prior to it.”

feelings” (GA I/5, 108). At the end of this genealogy of drives, its close connection to the concept of will becomes apparent.¹⁵

In the first part of the *System* we find a key definition of the term drive in § 3, in which Fichte defines it as an objective manifestation of the inner tendency of the I: “The posited tendency necessarily manifests itself in relation to the entire I as a drive” (GA I/5, 58). As Fichte contends here, we become conscious of our inner tendency when we look at the entire I – this emphasis on the entire I is important, as we will see later – in an objective manner and in describing it as a drive. Therefore, Fichte states that the task of § 3 is to “observe how the I becomes conscious of its own tendency toward absolute self-activity as such” (GA I/5, 54). This consciousness is possible only in a procedure of reflection, in which the drive functions as “a real, inner explanatory ground of an actual self-activity” (GA I/5, 55). Fichte’s concept of drive depends on a special kind of reflection, as Fichte teaches us in § 3.

However, in this context Fichte actually distinguishes between two kinds of reflection. Already in § 2 he mentions reflection, but this first reflection cannot be identified with the reflection from § 3. The first is a reflection upon the I as such. As soon as we not only posit ourselves as the I, but also become aware of the I, reflection upon the I is needed. On the one hand, we are the I, and on the other, we are also conscious of the I. In this way, being the I and consciousness of the I must coincide (GA I/5, 46).¹⁶ In § 2 Fichte hence focuses on the question of how the I as an intellect intuitively itself. To gain knowledge about this consciousness of the I, we need to become “conscious of consciousness” (GA I/5, 47). In § 3 he summarizes the results of this first kind of reflection in the following way:

In the previous section we proceeded by simply postulating a reflection upon the objective I, which is what was there under consideration; and we were undoubtedly entitled to postulate that the I is necessarily an intellect, and indeed an intellect that intuitively itself unconditionally. We, in our philosophizing, were mere spectators of a self-intuition [IV, 40] on the part of the original I. What we established was not something we ourselves had thought,

15 Fichte’s *System of Ethics* offers a most systematic presentation of his theory of drives, although Fichte discussed this concept already in his earlier works and lectures. Cf. Binkelmann 2006, 6.

16 As Fichte points out here, we also have to be aware of an important difference between reflecting upon a thing and reflecting upon the I: “The difference between a thing and the I (a rational being), which is entirely opposed to the former, is precisely this: a thing is merely supposed to be, without itself having the least knowledge of its own being. In the I, however, being and consciousness are supposed to coincide [...]” (GA I/5, 46).

but something the I had thought. The object of our reflection was itself a reflection. (GA I/5, 54)

In § 3 and in relation to the concept of drive, however, another reflection comes into play. The reason why Fichte introduces yet another reflection lies in the first reflection's inability to focus explicitly on the tendency of the I. This means that the first reflection only reflects upon the I as a faculty (*Vermögen*), but it does not consider specifically the tendency of the I – this is its dynamic character. As Fichte maintains, with the first reflection only “the consciousness of a mere power (*Vermögen*)” was established, “but certainly not the consciousness of a tendency or drive” (GA I/5, 54). If we want to reflect on this tendency in particular, we have to conceive of the I as a drive. In this way, then, the entire I appears as a drive,¹⁷ and thus, as a manifestation of the inner tendency of the I:

Insofar as it is not a subject and not an object but rather a subject-object (which means nothing more than an empty place for thinking), this entire I contains within itself a tendency to absolute self-activity, which, when it is thought of apart from the substance itself and [IV, 43] as the ground of the latter's activity, is a drive that drives the substance. (GA I/5, 56)

We do not immediately conceive of this drive as a feeling, but instead this drive results in a thought. What does this mean? Fichte defines this thought as a result of the drive and conceives it as a concept through which we determine ourselves. In this way, Fichte comes close to Kant's conceptual causality and his notion of autonomy. Drive expressed as a thought explains how we can determine ourselves freely through concepts.¹⁸ The concept he has in mind here is the concept of absolute self-activity through which we determine ourselves:

The content of the thought we have derived can therefore be briefly described as follows: we are required to think that we are supposed to determine ourselves consciously, purely and simply through concepts, indeed, in accordance with the concept of absolute self-activity; and this act of thinking is precisely that consciousness of our original tendency to absolute self-activity that we have been seeking. (GA I/5, 61)

As we see from this quotation, Fichte's main reason for introducing the concept of drive is to elucidate his conception of the tendency of the I toward self-activity.

¹⁷ For further interpretation of Fichte's conception of the entire I conceived as a drive, see Binkelmann 2006, 11 and Wood 2016, 118f.

¹⁸ According to Allen Wood, this thought is “the whole I,” meaning the unity of the subjective and objective aspects of the I and as such this thought is necessary to imagine how the I can be free and self-determined. Cf. Wood 2016, 119.

The concept of drive is suitable for this task because it embodies its dynamic aspect. Fichte emphasizes this very clearly already in the first part of the *System*, but in the second part the concept of drive starts to play an even greater role. But the starting point in the second part is similar to that of the first part:

If one originally thinks the I in an objective manner – and this is how it is found prior to all other types of consciousness –, then one can describe its determinacy only as a tendency or a drive, as has here been sufficiently demonstrated right from the start. The objective constitution of an I is by no means a being or subsistence; for that would make it the opposite of what it is; i.e., that would make it a thing. The being of the I is absolute activity and nothing but activity; but activity, taken objectively, is drive. (GA I/5, 105)

Also in this quotation, Fichte describes tendency as drive, when viewed objectively. Hence, drive is clearly bound with activity and represents the dynamic nature of the I. However, there is also a crucial difference when compared with the first part. Here, Fichte focuses on the experience of the particular, determined I and in this case we experience the drive as a feeling, as Fichte now argues. Therefore, the link between determinacy and feeling becomes central to his definition of drive: “Thus, if the I is originally posited with a drive as its objective determination, then it is necessarily posited as well with some feeling of this drive” (GA I/5, 105).

There is, of course, a significant difference whether Fichte claims that we conceive of the drive as a thought, as in the first part, or as a feeling, as in the second part. Fichte has two ways of characterizing the concept of drive: the drive for the entire I appears as thought, but when he deals with a determined and limited I, then he refers to the feeling of the drive.¹⁹ This means that we can either represent the tendency of the I in particular acts and speak about the feeling of the drive; or we can speak about the dynamicity of the entire I and then this tendency of the I results in a thought that can bring about the self-determination of the I.

However, as we can infer from Fichte’s second part, it is actually the second drive as feeling, which we experience in a most immediate way. At this stage we conceive of ourselves still as a part of nature, and hence the I appears as finite and determined. Consequently, this means that we have a tendency toward self-

¹⁹ Cf. Wood 2016, 154: “The two drives are fundamentally the same, but in ordinary consciousness they are differently experienced. The natural drive is experienced as longing, which aims at satisfaction—even at satisfaction solely for its own sake, or enjoyment. By contrast, the pure drive is experienced as an absolute demanding (Fordern) (SL 4:145). It appears not as a feeling but as a thought—the thought of a categorical imperative commanding us to an action or omission purely for its own sake.”

determination and freedom, although we are not free and self-determined to begin with. On the contrary, only tendency, and hence striving toward freedom, is characteristic of Fichte's dynamic conception of the activity of the I. An absolute freedom that the entire I as a drive describes is in reality never completely fulfilled. But in both cases, regardless of the particular differences, it is always this dynamicity of the I that Fichte wants to emphasize with the help of his concept of drive. Hence, the drive is always conceived of as an objective representation of the dynamicity of the I.

4 Drive as an intermediary between nature and freedom

Fichte's theory of drives is, however, much more complex and manifold than the sketch I have presented so far. In the second part of the *System* Fichte develops a whole system of drives corresponding to different levels of reflection and comprising different forms: from natural drives (*Naturtrieb*) on the one end of the spectrum to the pure drive (*der reine Trieb*) on the other. Each form in this system represents a specific manifestation of the tendency of the I at a certain stage of reflection. Because Fichte's system of drives comprises many different stages of the I, he can make use of this concept not only when demonstrating that we are a part of nature, but also when he wants to emphasize how we strive to freedom and independence from nature.

Fichte's system of drives is genealogically structured. At the lowest level, we experience natural drives through feeling, but at the highest level Fichte mentions the pure drive, which corresponds with the drive for the entire I, introduced already in the first part of the *System*. As a result of this broad span, the concept of drive can play an essential role in mediating between nature and freedom. This mediation becomes particularly apparent especially in § 8, when Fichte discusses the antinomy between activity and cognition, for whose resolution he refers to the concept of drive.

Fichte starts § 8 with two opposing statements – thesis and antithesis.²⁰ The thesis states how cognition is possible: "Thesis. A rational being has no cognition except as the result of a limitation of its activity" (GA I/5, 103). In the thesis, then, Fichte understands cognition as a result of the activity of the I; however, not as a result of the unlimited, but as a result of a limited activity. As I pointed out in the previous section, it is impossible to be conscious of the pure activity of

20 For the interpretation of § 8 see Schick 2015, 77–91 and Wood 2016, 143.

the I. Only when we look at the activity from an objective perspective, can we become conscious of this activity. In the thesis, Fichte restates this claim and explicates in more detail why the experience we have of our self-activity is not yet an absolute, but rather a limited activity:

I find myself only as free, and I do so only [IV, 103] in an actual perception of a determinate self-activity. [...] There is no consciousness whatsoever without consciousness of self-activity; this self-activity, however, cannot itself become an object of consciousness unless it is limited. (GA I/5, 103)

However, a problematic result of Fichte's thesis that determinate self-activity precedes cognition is that the determination and limitation of the activity is only possible when there is some object to which consciousness refers: This is what the antithesis claims and thereby opposes the thesis. In the antithesis, Fichte states that the self-activity must be a consequence of cognition and hence activity cannot be the starting point: "Antithesis. But self-activity does not pertain to a rational being as such, except in consequence of a cognition, at the very least in consequence of a cognition of something in the rational being itself" (GA I/5, 103). This cognition does not need to be an outcome of the objects outside of the I, but it must at least be an effect of something objective in rational beings themselves.

In this way, Fichte demonstrates that thesis and antithesis are caught in a circle: self-activity depends on cognition, states the antithesis, but the thesis states that cognition depends on self-activity. However, Fichte of course also proposes a resolution to this circularity by interpreting activity and cognition not as two wholly disconnected elements but by understanding them as two sides of the same coin. To show this, he refers to the concept of drive that serves as a link between cognition and activity. Why does the drive have this intermediary status? One reason, as Fichte argues, lies in the synthesizing ability of the drive.

The drive, as a manifestation of the tendency of the I toward self-activity, combines both aspects. First, it brings to expression the self-activity of the I; however, Fichte defines this activity as a determinate and limited activity and therefore dependent on something passive – on objects. This activity is thus not a pure or absolute activity of the I. Therefore, the antithesis can claim that this limited activity depends on cognition. But these two opposing aspects are combined, as Fichte wants to convince us, in the original feeling of a drive and thus this concept of drive expressed as feeling can function as a synthetic element:

This original feeling of a drive is precisely the synthetic element we described above. A drive is an activity that necessarily becomes a cognition in the I, and this cognition is not, as it were, an image or anything similar of the drive's activity; it is this activity itself, immediately presented. If this activity is posited, then the cognition of it is also posited immediately; and if this cognition is posited – with respect to its form, as a feeling – then the activity itself is posited. (GA I/5, 106)

Hence, when we are aware of the activity of the I in this immediate way, namely through feeling, there is no distinction between cognition and activity yet; instead, they are united and correlated. This experience of activity through the original feeling of a drive is the first stage in Fichte's genealogy of the I. At this stage, we already experience the striving of the I; however, at the same time, we conceive of ourselves as a part of nature, and not yet free. Fichte describes this form of the drive as a natural drive and connects it with the feeling of limitation:

[D]rive and feeling are precisely the manifestation of what is bounded and of the boundedness (*Begrenztheit*) within us. There is therefore an original, determinate system of drives and feelings. – According to what was said earlier, what is fixed and determined independently of freedom is called nature. This system of drives and feelings is thus to be thought of as nature; and the nature in question is to be thought of as our nature, inasmuch as consciousness of it imposes itself on us, and the substance in which this system is located is at the same time supposed to be the substance that thinks and wills freely, which we posit as ourselves. I myself am, in a certain respect, nature, notwithstanding the absolute character of my reason and my freedom; and this nature of mine is a drive. (GA I/5, 108)

But nonetheless, we cannot neglect the tendency of the I toward absolute self-activity. As the concept of drive shows, we experience this feeling of the drive precisely because of the striving of the I. Hence, the essence of this limited I still lies in its tendency to self-activity and freedom. Even when Fichte speaks about the natural drives, he also wants to emphasize the presence of this tendency toward self-activity and indeterminacy. The objective manifestation of this tendency as such, and not the expression of the tendency in particular acts, is what Fichte calls pure drive or higher drive. The concept of pure drive corresponds with his drive of the entire I, mentioned in the first part of the *System*. This drive is, as Fichte maintains, “drive for freedom, simply for freedom's sake” (GA I/5, 132). Through the process of development of reflection, we can gradually overcome our dependency on nature and emancipate ourselves. Again, this is possible because the essence of the I lies in the tendency to self-activity and absolute self-determination.

To reiterate: the concept of drive functions in Fichte's *System* as a synthetic link between nature and freedom because it brings to expression the determina-

cy of the I as well as the tendency to overcome its natural determination and to posit itself independently of nature. With the help of the drive, we can explain rational beings as a part of nature – they are embodied in a natural organic body – but at the same time as free in their striving to overcome nature. In this way the concept of drive unifies both aspects – activity and cognition – and mediates between nature and freedom, respectively.

5 From the concept of drive to the concept of will

So far, I have given a brief overview and presented my reading of Fichte's concept of will as well as of his theory of drives, but what is still missing is the discussion of the connection of the will and the drive. In this section, I will focus on this link between both concepts and point out the similarities between them as well as their differences. According to my interpretation, Fichte's concept of drive is of crucial importance for defining and explaining the concept of will. To substantiate this claim, I will compare Fichte's system of drives with his definition of the will from the third part of the *System*, especially from § 14, in which Fichte's discusses the will in particular.

Fichte starts this paragraph with the definition of the will as a free transition from indeterminacy to determinacy: "An act of willing [*Wollen*] is an absolutely free transition from indeterminacy to determinacy, accompanied by a consciousness of this transition" (GA I/5, 147). This definition corresponds with my discussion on his concept of willing in the second section. Moreover, we can find a similar definition of the act of willing that relies on the transition from what is subjective into what is objective also in the second part of the *System*:

The mere representation of an act of willing is the very representation we have just produced within ourselves: the representation of an absolute transition (accomplished by means of absolute self-activity) of what is subjective into what is objective, for this is precisely the universal form of all free willing. (GA I/5, 90)

What Fichte wants to point out with this definition of willing as an absolute transition from indeterminacy to determinacy (or from subjective to objective) is, in my view, the process of the self-determination of the will. Because of our tendency to self-activity, we strive to determine ourselves freely and independently of outer circumstances. At the same time, however, this also means that we are not self-determined to begin with, but that instead this happens in a process, in which we become conscious of our freedom and of our power (*Vermögen*)

of self-determination. Fichte calls this transition a “universal form of all free willing” (GA I/5, 90). Hence, his free will must be a result of this procedure.

But how then is Fichte's concept of will connected to his theory of drives? As we have observed in the previous sections, Fichte defines the will and the drive as an objective manifestation of the tendency of the I toward self-activity. According to this definition, we might assume that the meaning of these two concepts is very similar. But this cannot be true without exceptions because Fichte's concept of drive is not always and necessarily connected with freedom, as his concept of natural drive shows. Fichte's definition of willing as a transition from indeterminacy to determinacy, by contrast, must include the ability to freely determine itself. To this extent, it seems reasonable to assume that there must be a crucial difference between Fichte's willing and his conception of drive. This indeed is what Fichte points out in § 14:

The will (*Wille*) is neither the drive nor the longing nor the desiring. In the case of a drive there is a propensity and an inclination; in the case of desiring there is, in addition, consciousness of the object of the inclination, but instead of any determinacy of the active I there is only indeterminacy. Desiring wishes that its object would come to it, but it does not want to move either hand or foot to bring this about. Determinacy ensues as the result of an act of willing. (GA I/5, 147)

In this quotation Fichte mentions four different kinds of propensities: the will, drive, longing, and desiring. In the second part of the *System* it became clear that longing and desiring belong to different kinds of drives. Therefore, Fichte first gave us a general characterization of these drives: they describe a propensity or an inclination for something. Afterwards he also distinguished between two particular kinds of drives, namely longing and desiring. In both cases there is still an inclination present, but in the case of desiring we are also conscious of the object of inclination (GA I/5, 122), whereas this consciousness of the object is missing in the case of longing. With respect to longing, Fichte states that it “arises from reflection upon this drive” (GA I/5, 120). At first drive is not my product but a product of nature, though when I become conscious of the drive, I become aware that it is my drive and hence within my power.²¹

The crucial distinction that Fichte wants to highlight in the second part of the *System* as well as in the latter passage from § 14 is that there is a difference

²¹ “It [drive] is given, and it does not depend upon me in any way. Nevertheless, I become conscious of this drive, and what it brings about within consciousness is something that stands within my power; or, more precisely, the drive does not act efficaciously within consciousness, but it is I who act efficaciously or do not act efficaciously, in accordance with this drive” (GA I/5, 121).

between drives as inclinations and his concept of will. How do we have to describe this difference then? Fichte compares desiring with willing and claims that in the case of desiring the I is still undetermined, whereas willing brings about the determinacy. In Fichte's language this means that desiring is able to reflect upon the object of desires and to decide either to follow them or not, but it cannot freely determine objects. In the case of the will, however, the object is a result of the will. For this determinacy, the will must be able to determine itself freely and posit its own objects; hence, the will must have its own causality, namely causality of concepts. Therefore, Fichte defines the act of willing as an act of transition from indeterminacy to determinacy. What Fichte describes here is nothing other than Kant's conception of the autonomy of the will, even though Fichte uses the term self-activity to describe this self-determination of the I.

As we see from this analysis, even though there is an important distinction in his definition of the will and the drive, the transition from one stage to another, and hence also the transition from desiring to willing, must nevertheless be smooth. What Fichte points out are different stages of reflection, but there is no clear-cut distinction between drives and the will. Even when Fichte writes about natural drives, there is already some striving toward self-activity present; however, striving at this stage is still manifested as a part of nature. Yet at the highest level of reflection, when Fichte discusses the will, this tendency to self-activity becomes fully conscious of itself and hence able to determine itself. Therefore, Fichte's conception of the will is a special kind of drive, or as Fichte describes this, an original drive is also present in willing.

The transition from one stage to another is gradual and we see this confirmed by his distinction between two different kinds of desiring. When Fichte discusses desiring in the upper quotation, this desiring corresponds to the "lower power of desire" (*das niedere Begehungsvermögen*, GA I/5, 122), which he nevertheless already connects with freedom, namely with formal freedom. In § 14 this formal freedom is expressed in his concept of will as *Willkür*, namely as a power of choosing.²² In the case of formal freedom, the I already chooses freely and determines its objects, but this determination is not yet a result of the complete self-sufficiency of the will. As a result, the will functions as a *Willkür*, i.e. as the power of choosing. *Willkür* can reflect upon nature, but it is not fully independent from nature and it does not determine the objects freely and

²² For the discussion of formal freedom and *Willkür* see Neuhouser 1990, 143–166; Kosch 2013, 150–168; Wood 2016, 66–71.

for itself. Therefore, the *Willkür* is not the pure moral will, but only a neutral power of choice between different objects.

The higher power of desire (*das obere Begehrungsvermögen*, GA I/5, 126), however, corresponds with his definition of the will as a transition from indeterminacy to determinacy. The concept of will as *Wille* means the self-determined will, which Fichte links to the concept of material freedom.²³ When the I is also materially free it gives itself the object freely and only then can we speak about the total indeterminacy of the will to determine the object.²⁴ Fichte connects this higher power of desire with material freedom and hence with the ability for self-determination.²⁵ Hence, this concept of the power of desire is already what Fichte calls a free will in the upper quote.

Whether Fichte discusses formal or material freedom, or whether he refers to an earlier stage of his genealogy of the I, in all these cases he points out an original drive as an expression of our inner tendency of the I toward self-activity and freedom.²⁶ This tendency is the driving force of our I and as such it confirms that the essence of the I indeed lies in its endless activity which Fichte describes as tendency or striving. As a result of this characterization, his concept of will must also be a manifestation of this activity. Therefore, Fichte's concept of will cannot be conceived of as a static substance but is instead described as a dynamic concept that brings to expression this activity of the I, which is its inner tendency not only to activity as such, but also to self-activity.

23 Cf. Wood 2016, 147: "This is what Fichte calls 'the pure drive'; its object has to be freedom simply for freedom's sake. By acting on this drive, we achieve material freedom."

24 This ability to determine ourselves freely and hence posit our own objects Fichte calls an ethical drive, as a drive that is comprised of natural drive and the pure drive. If we want to act in the world and be efficacious, we need to have an impact on the natural world, and hence a synthesis between natural and pure drives is needed. Cf. Binkelmann 2006, 17.

25 Cf. Wood 2016, 147: "Transcendental consideration of the two acts of reflection brings to light another aspect of the distinction between formal and material freedom. We are formally free when we reflect on a natural drive and become aware of it as non-determining (SL 4:135). The first act of reflection on a drive makes us aware of our formal freedom, the indeterminacy of our will. The second makes us aware of a possibility of acting independently of our natural drives." See also Wood 2016, 174: "Formal freedom is what I am aware of as a rational being when I know that, within the range of free choices presented to me by my situation, I am able to resist any drive or desire, so that the way I act depends on my free choice alone. Material freedom is what I achieve through every action that obeys the moral law by following my conscientious conviction about my duty."

26 Cf. GA I/5, 125: "Are my drive as a natural being and my tendency as a pure spirit two different drives? No, from the transcendental point of view the two are one and the same original drive (*Urtrieb*), which constitutes my being, simply viewed from two different sides."

6 Conclusion

As my analysis has shown, what connects nature and freedom in Fichte's system is a concept of drive, as a drive to freedom. Already at the lowest stage, although we do not yet reflect upon ourselves, but experience ourselves as a part of nature through feeling, we are nevertheless more than just nature because of this drive to freedom. Afterwards, when we gradually become conscious of this tendency and gain the ability to reflect upon our natural drives, the consciousness of freedom arises as well. Only during this later stage does Fichte begin to address the concept of will, which is always linked to freedom and the ability to determine ourselves independently of nature. However, according to this view, material freedom as the highest expression of our drive to freedom, is not something that we really find accomplished; rather, this material freedom as an absolute transition from indeterminacy to self-determination remains for us only an ideal, namely only a tendency for freedom, simply for freedom's sake. As a result of the structure of Fichte's system his ethics brings to expression a dynamic process, in which his concept of will is also a part of this process and must therefore be described in dynamic terms.

Therefore, the concept of drive plays a crucial role in transforming Fichte's moral philosophy into a system of ethics. Fichte's system of ethics is not a system because it unifies various unconnected elements and combines them into a system. On the contrary, Fichte's system is a dynamic system that demonstrates how different elements in this system evolve from each other. Dynamicity is a result of his notion of tendency – and hence, when taken in an objective manner, of the notion of the drive. As a result, drive functions as an intermediary link between nature and freedom. Therefore, the domains of nature and of freedom are in Fichte's system not just opposed to each other; instead they are also essentially connected and interdependent. In this sense, the sensible and a posteriori part of nature, which appear in Kant's philosophy always as something foreign to his idealism, are in Fichte's philosophy conceived as an integral part of his system of ethics. However, due to this original drive to self-activity and freedom, this system is in its essence still a system of freedom and not a system of nature.

Bibliography

- Beiser, Frederick C. (2002): *German Idealism. The Struggle against Subjectivism*. Cambridge: Harvard University Press.
- Binkelmann, Christoph (2006): "Phänomenologie der Freiheit. Die Trieblehre Fichtes im System der Sittenlehre von 1798." In: *Fichte-Studien* 27, pp. 5–21.

- Breazeale, Daniel (1996): "The Theory of Practice and the Practice of Theory: Fichte and the 'Primacy of Practical Reason.'" In: *International Philosophical Quarterly* 36, pp. 47–64.
- Breazeale, Daniel (2006): "Die systematischen Funktionen des Praktischen bei Fichte und dessen systematische Vieldeutigkeit." In: Hans Georg von Manz/Günter Zöller (eds.): *Fichtes praktische Philosophie. Eine systematische Einführung*. Hildesheim: Olms, pp. 39–72.
- Breazeale, Daniel (2012): "In Defense of Fichte's Account of Ethical Deliberation." In: *Archiv für Geschichte der Philosophie* 94, pp. 178–207.
- Fichte, Johann Gottlieb (2005): *The System of Ethics According to the Principles of the Wissenschaftslehre* (1798). Translated and edited by Daniel Breazeale and Günter Zöller. Cambridge: Cambridge University Press. [Paging follows Fichte's Gesamtausgabe, div. I, vol. 5, abbreviated as GA I/5]
- Kant, Immanuel (1996): *Practical Philosophy*. Translated and edited by Mary J. Gregor. Cambridge: Cambridge University Press. [References are by volume and page of the Academy Edition]
- Kosch, Michelle (2013): "Formal Freedom in Fichte's System of Ethics." In: *Internationales Jahrbuch des Deutschen Idealismus/International Yearbook of German Idealism* 9, pp. 150–168.
- Kosch, Michelle (2018): *Fichte's Ethics*. Oxford: Oxford University Press.
- Metz, Wilhelm (2006): "Freiheit und Reflexion in Fichtes Sittenlehre von 1798." In: *Fichte-Studien* 27, pp. 23–35.
- Neuhouser, Frederick (1990): *Fichte's Theory of Subjectivity*. Cambridge: Cambridge University Press.
- Owen, Ware (2010): "Fichte's Voluntarism." In: *European Journal of Philosophy* 18, No. 2, pp. 262–282.
- Owen, Ware (2015): "Agency and Evil in Fichte's Ethics." In: *Philosophers' Imprint* 15, No. 11, pp. 1–21.
- Owen, Ware (2018): "Fichte's Normative Ethics: Deontological or Teleological?" In: *Mind* 127, Issue 506, pp. 565–584.
- Quante, Michael (2015): "Alles geht aus vom Handeln, und vom Handeln des Ich – ein analytischer Kommentar zu §§ 4–7 von Fichtes *Sittenlehre*." In: Merle, Jean Christophe/Schmidt, Andreas (eds.): *Fichtes System der Sittenlehre. Ein Kooperativer Kommentar*. Frankfurt a.M.: Vittorio Klostermann, pp. 57–73.
- Schick, Friederike (2015): "Vermittlungen zwischen Natur und Freiheit – der Naturtrieb in § 8 des *Systems der Sittenlehre*." In: Merle, Jean Christophe/Schmidt, Andreas (eds.): *Fichtes System der Sittenlehre. Ein Kooperativer Kommentar*. Frankfurt a.M.: Vittorio Klostermann, pp. 75–92.
- Wood, Allen (2016): *Fichte's ethical thought*. Oxford: Oxford University Press.
- Zöller, Günter (1998): *Fichte's Transcendental Philosophy. The Original Duplicity of Intelligence and Will*. Cambridge: Cambridge University Press.

Ansgar Lyssy

Reality as Resistance: The Concept of the Will in Bouterwek's *Idea of an Apodictic* (1799)

Abstract: Friedrich Bouterwek's *Idea of an Apodictic* (1799) is a lengthy defense of the core claims of Kant's first *Critique* against different types of skeptical criticism. Bouterwek argues that Kant's presupposition of knowledge is vulnerable to skeptical attacks and he suggests shifting the apodictic and hence foundational commitment to reality from the domain of theoretical knowledge towards that of practical agency. Reality is not represented in our knowledge, but instead experienced directly as resistance against our will. The will is not a mere idea within a transcendental framework of theoretical speculation, but a proper force, a real entity that we can experience directly and that dissolves the abstract distinction between reality and thought.

1 Introduction

In the aftermath of the publication of Kant's groundbreaking *Critique of Pure Reason*, transcendental philosophy had quickly become both a major topic and method of German philosophy for the next few decades; even those who strongly disagreed with Kant simply could not ignore him. One of the lesser-known philosophers of this first wave of Kantians was Friedrich Ludewik Bouterwek (1766–1828, professor for philosophy in Göttingen since 1802). He had quickly written several books and a couple of short papers picking up on very different Kantian ideas. Arguably his main work is the *Idea of an Apodictic* (*Idee einer Apodiktik*, first published in 1799, henceforth cited simply as *Apodictic*, followed by volume number and page number¹). It contains a number of astute insights into the shortcomings of Kant's new philosophy, but it is also written rather hastily and shoddily, leaving significant gaps in its argumentation. However, a number of highly original ideas stand out, especially his conception

¹ Citation follows Bouterwek 1799/2018. The page numbers of the first edition are also indicated by means of marginal numbers in the 2018 edition, thus all page numbers given here refer to the first edition.

of the will. Since Bouterwek's work is rather unknown,² I will hereby first provide a broader introductory overview of his work for the sake of providing some context and making an unknown, but interesting thinker slightly more accessible. I will then attempt to sketch his main perspective on transcendental philosophy as it is discussed in the *Apodictic*, which will serve as a basis for my account of his conception of the will. Finally, I will outline both the strengths and weaknesses of this conception.

However, what good can come from examining such an admittedly obscure philosopher? Bouterwek's conception of the will is clearly sufficiently original in itself to merit some discussion, if only for the fact that he anticipates and pre-dates later positions (see section 5 of this paper). But beyond that, Bouterwek is also a small, but not insignificant figure within the *constellation* of classical German philosophy, i. e. the space of dialogue in which major and minor thinkers rapidly and interrelatedly respond to similar or identical problems (on the topic of constellations, cf. Henrich 1991). For example, the *Idea of an Apodictic* was reviewed by Reinhold, his *Principles of Speculative Philosophy* (*Anfangsgründe der spekulativen Philosophie*, 1800) by Hegel, and his *Sketch of Academic Lectures on Aesthetics* (*Grundriss akademischer Vorlesungen über die Ästhetik*, 1797) by A. W. Schlegel. Indeed, this is just a salient list of examples of his books being reviewed by the leading thinkers of his time.³ The famous scientist Johann Friedrich Blumenbach wrote Bouterwek's obituary (Blumenbach 1828, 3–10), and his monumental history of literature from the Renaissance onwards is still considered to be one of the foundational attempts at writing a comparative history of literature.⁴ He most likely has had a broader influence as well that remains to be traced, as he was a close friend of Jacobi (beginning in 1799) and Schopenhauer's teacher in Göttingen. It seems very likely that here we find a neglected conceptual resource into which at least Schopenhauer tapped to build his own philosophy of the will.

² The only comprehensive accounts of his thought and development are Struck 1912 and Senne 1972. Shorter introductions can be found in Adickes 1896 and Zinkstok 2010. A good account of the *Apodictic* can also be found in Dierse 1999.

³ For a full list of reviews of Bouterwek's works, see Senne 1972, 650 ff.

⁴ See Bouterwek 1801-. On Bouterwek's role in comparative literature studies and the history of the German *Literaturwissenschaft*, see eg. Zymner/Hölter 2013, 350; see also Melo Araújo 2012.

2 Friedrich Bouterwek

Despite being one of the earlier Kantians and an influential academic around the turn of the 18th century, Bouterwek's writings are little known today. The *Idea of an Apodictic* (1799) is his most accomplished work, a lengthy defense of the core claims of Kant's first *Critique* against different types of skeptical criticism.

Yet nonetheless, despite the fact that he was an established and well-connected intellectual and a very prolific writer whose work touched on almost all subjects of the philosophical debate in the later 18th century philosophy, he is largely forgotten. This may be partially due to his propagation of a specific type of realism, which will be discussed below and which went at odds with the emerging strand of German Idealism: Bouterwek emphasized the necessity of including a robust concept of reality into philosophy and the need to conceive of reality as something that is entirely independent from mind or cognition. He also may have overextended himself and his ambition may have been greater than his actual philosophical and scientific capacities – he deals not only with philosophical questions, but also writes about history, legal issues, literature, mineralogy, and classics. He confesses to being unable to decide which field of study to give priority.⁵

Today, Bouterwek may be best known as a *Popularphilosoph*, one of the main figures within the movement to liberate philosophy from a rigid, overtly abstract and academic framework and make it accessible to the public and applicable to everyday life. This movement is prevalent in his writings in two ways. First, many of his philosophical writings use a rather literary style, even when dealing with highly abstract topics. His *Paullus Septimius* (1799), for example, uses a theatrical form of dialogue to discuss and elaborate Kant's first *Critique*, which is, with intense emotion, described as a quasi-mystical awakening.⁶ His texts are often defined by an abundant use of rhetorical tropes that are used for frequent emphasis and sometimes sound overly oratorical, sometimes even preachy. Here, Bouterwek belongs stylistically more to the predecessors of the *Sturm und Drang*-movement. Second, Bouterwek emphasizes that philosophy is not a merely abstract exercise in developing knowledge, but something that

5 “Könnte ich mich nur, wie ihr Gelehrten, in irgendeine Wissenschaft verlieben! Aber dazu geb' ich alle Hoffnung auf. Mir bleibt nichts übrig, als mit ihnen allen ein Schmetterlingsleben zu führen: und zu gutem Glück ist der Kreis, in dem ich umherflattern kann, so groß, daß die erste, wenn die Reihe wieder an sie kommt, mir wieder ganz neu scheinen kann.” Letter to Johann Friedrich Pfaff, 1. Nov. 1789, in: Pfaff 1853, 206 f. The spelling of all citations has been carefully modernized.

6 On the *Paullus Septimius* see Lyssy 2014.

ideally provides peace of mind, that helps to develop a more humane and refined character, and that provides practical orientation in a time defined by profound social and intellectual challenges to the *status quo*. Theory and practice need to be related to each other, thought and feeling need to be consolidated. Some texts aim at readers far beyond academia, especially in a journal he edited called *Neue Vesta: Kleine Schriften zur Philosophie des Lebens und zur Beförderung der Häuslichen Humanität* (11 vols., 1803–1810). Here he and other contributors wrote populist texts that resemble entries in a modern lifestyle magazine or a self-help book, full of sentiment and even *Schwärmerei*.

But it would be too easy to dismiss Bouterwek's genuine philosophical writings as shallow because of his populist outreach. He lacks neither ambition nor an understanding of the problems of constructing a renewed form of 'first philosophy' after Kant. In fact, he strives to be part of the wave of 'foundational philosophy' (*Fundamentalphilosophie*) that swept through Germany in the aftermath of the first *Critique*, looking for the very first principles of thought and reason in general. Thus, the aims and intentions of the *Apodictic* are comparable to the works of Reinhold, Jacobi, Fichte, and Schulze (the latter under his *Aenesidemus* pseudonym) as there is some conceptual overlap between specific ideas of the *Apodictic* and those four philosophers. Some influence will probably be due to a similar reception of earlier thinkers such as Locke, Leibniz, Hume, and, of course, Kant. While Bouterwek often cites or at least mentions his sources, he never mentions, for unknown reasons, Jacobi, who he surely was acquainted with.⁷ I will avoid any discussion of possible influences here for the sake of brevity.

Kant himself was supportive of Bouterwek's works.⁸ When in 1793, six years before the publication of the *Apodictic*, Bouterwek sent his lecture manuscripts concerning Kant's philosophy to Kant himself, he received praise for his grasp of the matter and his capacity to make it more accessible to the public.⁹ When Bouterwek later articulated some criticism in his review of Kant's *Doctrine of Right*, Kant took these arguments serious, especially Bouterwek's critique of Kant's conception of marriage law,¹⁰ and indeed changed some parts of his text to conform to this criticism. Bouterwek clearly sees himself as a proper Kantian and hopes to

⁷ Adickes suggests that the contact with Jacobi's thought happened during the writing of the *Apodictic* and is visible in the latter half. See Adickes 1896, 411.

⁸ Kant cites a phrase from a poem attributed to Bouterwek in his *Toward Perpetual Peace* (see AA VIII, 367), but it rather seems to stem from Gottfried August Bürger's book *Der Vogel Urselfbst, seine Recensenten und der Genius*, which was published in 1789.

⁹ See Kants letter to Bouterwek from May 7, 1793. In: AA XI, 431.

¹⁰ See Ludwig 1988, esp. 138 ff.

improve on the first *Critique* so as to make it more defensible against skepticism, which he conceives to be the major weakness. But he is also interested in Kant's other projects, as seen in the *Five Letters on Cosmopolitanism* (*Fünf kosmopolitische Briefe*, 1794) where Bouterwek responds to (among others) Kant's *Idea of a Universal History*; and in *Morriston or who can give commands?* (*Morriston, oder: Wer hat zu befehlen?*, 1798) he treats the perceived shortcomings of Kant's *Doctrine of Law*.

Later, however, Bouterwek abandons Kantianism in favor of Jacobi's *Glaubensphilosophie* (cf. his *Lehrbuch der Philosophischen Vorkenntnisse*, 1810), with whom he even maintains a long personal friendship. The final two decades of his life are dedicated mostly to his ambitious history of literature and he writes much less on philosophical topics.

The *Idea of an Apodictic*¹¹ is, as mentioned above, Bouterwek's major work in theoretical philosophy. In this work, we find an original account of the will, aimed at overcoming the alleged deficiencies of Hume's, Leibniz's and especially Kant's conceptions. It merits an inquiry since, as we will discuss later, it also has some major ideas in common with some contemporary accounts of experience and the will. But before I discuss the theory of the will, I want to outline the general project of the *Apodictic*. Next, I will outline his take on Kant's failings and other attempts to ground philosophy, and afterwards I will present and finally discuss his own solution.

3 The *Apodictic*

Very broadly speaking, Bouterwek is endeavoring to answer the questions raised by Kant and aiming at avoiding the shortcomings of the *Critique of Pure Reason* (henceforth: *CPR*). Bouterwek picks up the question of how proper knowledge is possible. He reads Kant's *CPR* as a theory of propositional knowledge that strives for an apodictic foundation of epistemology instead of a critique of metaphysics. The promising angle developed by Kant is to approach philosophical problems from the point of view of 'meta-philosophy,' i.e. every philosophical idea about *x* must also concurrently state the validity of its own reasoning regarding *x*. This is also the general method Bouterwek adopts in the *Apodictic*, where he raises a second question related to the possibility of proper knowledge: How is philosophy possible?

11 The first part of the *Apodictic* is now available in a new edition (see also Bouterwek 2018).

Just as several other contemporaries, Bouterwek finds that Kant's philosophy, especially the deduction of the categories, is vulnerable to skeptical attacks. Here, he follows Schulze's *Aenesidemus* (1792) and concedes that Kant has little to offer as a counterargument. He argues that Kant's premise is the assumption that we do, in fact, have proper knowledge. This knowledge is made possible by the categories of reason that are grounded in transcendental apperception or the unity of consciousness. The validity of the categories is supposed to enable objective knowledge, and if their validity is questioned, then so is the possibility of all knowledge that has some claim to intersubjective, non-negotiable truth. Transcendental apperception, as a merely formal element of cognition, cannot serve as the unshakeable ground of all proper, that is, non-contingent or a priori cognition on which the objectivity of all empirical cognition ultimately depends.

But this presupposition that we have actual, objective knowledge instead of merely presumed knowledge is vulnerable to skeptical attacks because it entails the claim that this knowledge has objective validity. While Kant claims to derive objectivity from the categories, he already presupposes that this claim to objectivity is fulfilled in his assumption that we do have, in fact, objective knowledge. The skeptical argument thus is very simple: Kant argues that, since we know something, we must concede certain necessary structures of cognition, namely the categories. However, the skeptic could question this assumption: how do we know that, in fact, we do actually know something? Consequently, Kant is trapped in a circle: If Kant wants to plausibly assume that we do have some proper cognition, he needs to establish the possibility of this cognition first; nevertheless, to establish the possibility of this cognition, he assumes that we do have, in fact, this cognition.

Instead, what we need first, according to Bouterwek, is a proper proof that at least some of our knowledge has objective validity; only then can an inquiry into categories and other transcendental conditions be successful. This is what he calls the "ground of all proofs" (*Apodictic* I, 8). This foundational principle that needs to be found should guarantee the objectivity of all truths found through reason.

Part of Bouterwek's reasoning can be summarized as follows. All concepts or conceptual judgments presuppose a claim to objectivity, but as all concepts are *general* in nature, they also fail to capture the *particularity* of reality. In virtue of this, concepts *alone* cannot sufficiently guarantee their own objectivity to anything real within their own conceptual framework. Thus, all conceptual approaches to an apodictic foundation of knowledge will have to refer to something outside of them.

Bouterwek also emphasizes the distinction between theory and practice. Within theory, one needs to distinguish thought (*Denken*) on the one hand,

which develops the logical connection between concepts, from knowledge (*Wissen*) or cognition (*Erkenntnis*) on the other. Only knowledge and cognition can claim objective validity; and only by assuming that objective validity of knowledge is possible can we form proper theories and insights by connecting them in our thought. Mere thinking is empty, for it can be disconnected from all reality, while cognition and knowledge cannot. Both, however, need some direct or indirect grounding in the experience of reality, an experience that is not identical to the cognition of reality. This is where another criticism of Kant comes into play: Kant's transcendental approach to experience prevents it from corresponding to reality since it only takes place in the phenomenal world.

There is a non-conceptual element to experience (the will, to which we will soon turn) that may not serve as an immediate ground of knowledge, but it can, as Bouterwek contends, serve as a ground of certainty inasmuch as it does not fall into the same trap of generality as concepts do. Once a basic certainty about the existence of reality in a very broad sense is established through non-conceptual experience, objective validity of cognition and knowledge can be salvaged. This is what Bouterwek calls his "Practical Realism" (*Apodictic* I, xxiii.) – reality *precedes* knowledge and can be experienced *directly*; acknowledgment of such a reality can and should be seen as the condition of the possibility of all objective knowledge. While both Bouterwek and Kant strive to find the grounds of objective knowledge, Kant argues that formal elements of thought (i.e. the categories) alone are not only sufficient criteria; they are also the only criteria for objectivity that we can find. Bouterwek, quite to the contrary, argues that besides the categories there is also a type of certainty that comes with specific ideas and experiences and that is not a formal element of thought, but a material aspect that is common to the broad spectrum of our experiences.

Bouterwek uses these distinctions to structure the *Apodictic*, which is divided into two volumes, each containing two books. The first book (*Logical Apodictic*) deals with mere thought, whereas the second one (*Transcendental Apodictic*) with cognition or knowledge, i.e. thoughts with an objective claim. In both of these first books, he tries to highlight the failings of Kant (and others who have tried to establish a *prima philosophia*) to ground knowledge and to justify our claims to truth and objectivity. Only in the third book (*Practical Apodictic*) is a proper solution articulated. The fourth book (*Philosophical Syntax – or On the Arrangement of Ideas in the Spirit of the Apodictic*) formulates some consequences of his approach.

Bouterwek thus suggests shifting the apodictic and hence foundational commitment to objective validity from the domain of theoretical knowledge towards that of practical agency – which amounts, in short, to the idea that reality is not, in the first place, represented in our knowledge, but originally experienced di-

rectly as resistance against our will. A proper foundation of cognition would have to be rather different and could not be immanent to cognition. Put differently, cognition needs to be grounded in *something* that is not immediately cognition itself; reason also needs to be grounded in *something* that is not reason itself. Somewhat similarly to Fichte, he argues that the formal distinction between subject and object must be grounded in the subject; but he asserts, against Fichte, that the content of the object cannot be derived from the subject alone (cf. *Apodictic* II, 382ff.). Just as Jacobi had claimed shortly before him, the experiences of one's own reality and of an external reality cannot be separated.

Central to Bouterwek's project is therefore the notion that the connection between my own act of willing and an external reality is experienced without any concepts involved: the content of thought cannot be determined by the will on its own and thus it cannot originate from the subject alone. By means of the will, we are thereby given a form of knowledge that is distinct from knowledge that is gained through thought. Bouterwek shifts the presupposition that is supposed to serve as the starting point of all philosophy from the claim of our possession of knowledge towards the claim that, in fact, something does exist – and this ontological claim is much less susceptible to skeptical attacks than the claim of knowledge. The will is not a mere idea within a transcendental framework of theoretical speculation, but a proper force, to wit: a real entity that we can experience directly and without any prior knowledge, even without any primordial concepts. This foundational repositioning of the will is supposed to 'dissolve' the abstract distinction between reality and thought in our experience. Metaphysics, epistemology, and practical philosophy can, Bouterwek hopes, all be grounded in this primordial act of experiencing reality through the will.

4 Elusive Reality

Let me now briefly sketch Bouterwek's criticism of the attempts by Kant and other philosophers to find an apodictic ground for thought and reason in general. This philosophical critique is developed in broad strokes in the first two books of the *Apodictic*.

Since knowledge or cognition cannot be grounded in itself, theoretical philosophy, defined as the philosophical inquiry into the nature and scope of proper knowledge, is insufficient. It requires a search for an 'elementary principle' (*Elementarprinzip*) that could ground the system of thought and knowledge against skeptical attacks concerning the objectivity of cognition. It serves as an unshakable starting point from which all other philosophical assumptions can derive their validity. Without being able to go into all the details here, Bouterwek dis-

cusses a couple of promising solutions for such an elementary principle, which he associates with the philosophy of Descartes, Locke, Leibniz, and Kant. Here he finds various proposals regarding some ‘elementary concepts’ such as thought, reality (*Dasein*) and the “I” that have allegedly been used to ground proper knowledge. But none of them, he submits, can serve as an apodictic ground (*apodiktische Begründung*) or even a proper starting point (*Anfang*) for knowledge since they are all concepts of reason, yet, as we have just seen, reason cannot ground itself in itself. It hereby requires a proper grounding of arguments in reality (*Realgrund*) that cannot be logically deduced. In doing so, Bouterwek picks up Crusius’ distinction between *Realgrund* and *Idealgrund*, which roughly mirrors the scholastic’s distinction between the *ratio essendi* (the reason why something is) and the *ratio cognoscendi* (the reason why we know something). This distinction also appears, with slightly different emphases in each, in Kant, Platner, and Feder, the latter being Bouterwek’s predecessor in Göttingen. While Bouterwek does not offer a concise definition of these crucial concepts, the *Idealgrund* is the logical reason due to which we hold something to be true whereas the *Realgrund* is the material reason for the truth of our thought.

According to Bouterwek’s model, axiomatic systems of philosophy are unable to provide proper grounding since all they can provide are logical reasons (*Idealgründe*). This includes such systems as developed by Leibniz and Wolff; but Bouterwek also counts Kant among the axiomatic thinkers, as he also includes many postulates and unprovable principles (*Grundsätze*) in his philosophy (cf. *Apodictic* I, 118f.). Axiomatic systems may provide valid conclusions by means of syllogistic reasoning, but they need to assume that something exists in the first place that can ground the truth claims of all judgments involved in any syllogism; and no existence can be proved syllogistically.

Even an external absolute cannot serve as this elementary principle, since, *pace* Descartes, it cannot be proven directly through thinking. Even more so, the absolute cannot be sufficiently captured through mere concepts, since, again, concepts cannot provide their relation to reality itself.

Bouterwek therefore contends that the core problem of transcendental philosophy is a similar restriction of knowledge to the realm of appearances, thereby cutting it off from reality and preventing it from finding any actual ground. The proper conditions of knowledge are: experience and thought, as Kant had argued, but also reality itself, which Kant had ignored. This neglect on the part of Kant is not satisfying inasmuch as it fails to answer the question: *What is it that appears?* It is worth noting that in so criticizing Kant, Bouterwek apparently adheres to the “two-worlds” reading of Kant’s idealism and he implicitly rejects the interpretation of the thing in itself as a mere limiting concept. Appearance is appearance *of something* and not merely nothing. While we may not be

able to conceive of this something properly, we have to integrate some principle that ‘reaches out’ to it into our thinking.

Bouterwek argues, directed against Kant, that an idealist, even a transcendental idealist, would assume that thinking can ground itself by confirming to some internal criteria that provide objectivity. He claims that transcendental philosophy can thereby only unveil the subjective grounds of knowledge, but fails to establish the objectivity it is aimed at. Instead of conceiving the absolute as proper reality, transcendental philosophy demotes it to the status of a regulative ideal – and as such, it cannot refute any skeptical attacks on objectivity. Consequently, idealism entails the assumption that both thought and knowledge can contain their own elementary principles. Yet since knowledge can only be differentiated from mere thinking if it refers to things ‘outside’ of us, it thus needs some external grounding to guarantee its objectivity. Knowledge cannot establish its own objectivity from within and it can never arise from mere thinking alone. However, because the idealist fails to provide such an elementary principle, he or she also fails to properly distinguish thought from knowledge. The idealist sells us this assumed independence of thought as freedom, while it rather is a free-floating train of thought detached from reality. Bouterwek writes:

Natural human understanding recognizes this first determination of the logical forms by the mind itself [i.e. in this context: the transcendental subject] as an empty determination when it apodictically opposes thinking to knowledge. Naked thinking has little value for it; for fools and the wise think: but whether and to what extent there is truth in what we think – this is what we would like to know; and the logical factum of thinking cannot give us any information about it because this factum itself logically only has a scientific value insofar as we know that we think. (*Apodictic* I, 37f.)¹²

This grounding cannot be given through logic or any other attempt to provide a proper grounding for knowledge through the formal structure of judgments since judgments presuppose the fact that they are thought or conceived by the subject. Otherwise, what’s the point of analyzing judgments? Logic cannot explain its own dependency on thinking and consciousness (cf. *Apodictic* I, 49 ff.).

12 “Diese erste Bestimmung der logischen Formen durch den Verstand selbst (i.e. transcendental subject) erkennt denn auch der natürliche Menschenverstand als eine an sich leere Bestimmung an, wenn er das Denken dem Wissen apodiktisch entgegen stellt. Das nackte Denken hat für ihn wenig Wert; denn Toren und Weise denken: aber ob und in wie- / 38 / fern Wahrheit in dem ist, was wir denken, das ist es was wir gern wissen möchten; und darüber kann uns das logische Factum des Denkens gar keine Aufschlüsse geben, weil dieses Factum selbst logisch nur in sofern einen wissenschaftlichen Wert hat, als wir wissen dass wir denken.”

Thought itself, or the “I think” that plays such an important role in both Descartes, Kant and Fichte, also cannot be the ground of all knowledge, it serves as a proper ground for logical thinking. Instead of being a principle that guarantees the objective reality of the content of our knowledge (*Realprinzip*), it rather serves to ground the unity of all propositions if we abstract from their claim of validity and treat propositions as mere structures of thought (*Idealprinzip*). Both the logical and transcendental “I” are “nothing” (*Apodictic* I, 396), i.e. they contain no reality whatsoever.

Bouterwek rejects the Cartesian approach of using self-consciousness as the starting point of philosophy and our quest for the nature of reality; and he also rejects Fichte’s idea that objectivity has to lead back to an original self-positing of the “I” (cf. *Apodictic* I, 17 f., 168 f.). Bouterwek claims that the subject is neither prior nor posterior to the object. Therefore it is not in itself *given a priori*, but rather *discovered*:

The subject does not theoretically *posit* itself; it discovers itself like it finds an object. It discovers itself with the object as reality by doubling reality through absolute reflection. It is not a product of the opposing reality [*der entgegengesetzten Realität*] and the latter is not its product. It finds its being in the first place, independent of itself and independent of any other reality, as long as one is speaking of mere being, not of the *kind* of being. It has to be found in this way without being able to fathom theoretically why. (*Apodictic* I, 234)¹³

Both Kant and Fichte postulate the primacy of the subject, but Bouterwek claims that they fail to understand that without an object, a subject is not conceivable or intelligible (cf. *Apodictic* I, 237 f.). Nevertheless, Bouterwek accepts Kant’s proposal that all thought is unified in consciousness through its form, namely the “I think that *p*” that is enabled through transcendental apperception. But he disagrees that this can serve as a ground for objectivity or truth. Since all the content of cognition is made possible through transcendental apperception, Kant’s deduction ultimately leads to a circle: The determination of cognition is grounded in nothing but the fact that I think of it, which holds for all philosophical claims as well (cf. *Apodictic* I, 112f.). The “I” of the transcendental apperception still serves as an ideal principle for theoretical philosophy, but only as a formal prin-

¹³ “Das Subjekt *setzt* theoretisch nicht sich selbst; es entdeckt sich selbst; wie es ein Objekt findet. Es entdeckt sich mit dem Objekt als Realität, indem es durch absolute Reflexion die Realität verdoppelt. Es ist kein Produkt der entgegengesetzten Realität und diese ist nicht sein Produkt. Es findet sein Dasein überhaupt, unabhängig von sich und unabhängig von jeder anderen Realität, sofern von bloßem Dasein, nicht von der *Art* des Daseins, die Rede ist. Es muss sich so finden, ohne theoretisch ergründen zu können, warum.”

ciple, the unity of cognition. Again, it cannot be the elementary principle that Bouterwek is looking for, as it cannot ground any proper objectivity.

Bouterwek argues that Kant has denigrated the principle of absolute reality to a mere regulative principle and thereby stripped it from all predicates. Since Kant has refused to develop a robust concept of reality, all he has left is subjective experience that is ultimately ungrounded. The attempts of theoretical philosophy to construct a realism rather lead towards a “negative Spinozism” (for example, *Apodictic* I, 392f., 399): By conceiving the thing in itself as merely “*x*,” Kant has stripped it of all predicates and thereby also stripped reality of its claim to be absolute and delegated reality to the realm of appearances. While Spinoza has proposed reality as an absolute totality of predicates and predictability, for Kant it is, according to Bouterwek, just the opposite: namely, the absolute absence of such. This is why he designates this position “negative Spinozism,” with which Bouterwek concludes the *Transcendental Apodictic*: proper, mind-independent reality is to be conceived as an absolute absence of predicates. As such, it cannot serve as the foundational principle of philosophy.

Although Kant’s transcendental philosophy correctly conceives of reality as a regulative principle (e.g., in form of the three postulates of reason: the soul, the world, and God), it fails to understand how such a principle could be derived or grounded. The second part of the *Apodictic* that deals with the scope and reach of transcendental philosophy contains such a harsh criticism of Kant: He fails to ground objective knowledge, but instead rather proposes something that pretends to be knowledge. Within transcendental philosophy, metaphysical knowledge is nothing but absolute ignorance (cf. *Apodictic* I, 382). Bouterwek concludes that neither thought nor knowledge have yet been supported by such an elementary principle; this seems to be impossible, since they cannot encompass the reality they are referring to within their own framework.

5 Practical Philosophy and the Principle of Reality

After this criticism directed against Descartes, Locke, Leibniz, Kant and Fichte, Bouterwek attempts to solve the problems he sketched before. The foundation of knowledge cannot rest on the way the subject relates to an object in the act of cognition, at least not without some non-conceptual account of the object itself as well. Theoretical philosophy cannot provide an apodictic foundation for knowing since the “I think” cannot be grounded from a merely theoretical perspective. Moreover, it cannot sufficiently serve as a ground for objectivity itself

since it cannot transcend the realm of thoughts into the realm of reality. But, Bouterwek suggests, the focus on cognition, knowledge, and the foundational capacities of theoretical philosophy might simply be too narrow. Maybe practical philosophy can help? Given that skepticism has been conceived as a merely theoretical position without a practical dimension, practical ideas are not in the same way vulnerable to skeptical attacks.

Bouterwek begins by outlining the parallels between theory and practice and their respective philosophy. Theoretical philosophy is based on the elementary concepts of knowledge, consciousness, and necessity. The “I think” will serve as the unsurpassable, yet insufficient foundation that is not yet a real foundation of all theoretical philosophy, but merely its ideal basis (*Idealgrund*). Practical philosophy, on the other hand, is based on the elementary concepts of the will (*das Wollen*), conscience, and freedom. Its ideal foundation is not the merely contemplative “I think,” but rather the active and transformative “I want.” This will serve, Bouterwek continues, as the real foundation (*Realgrund*) of not only practical, but also of theoretical philosophy – not as a proposition, but rather as an underlying experience that precedes all propositions. The will is immediately certain and cannot be reasonably doubted (*Apodictic* II, 21f.) – it is as certain as being itself (*Apodictic* II, 30).

It is an essential argument in Kant’s transcendental deduction that the “I think” must be able to accompany any judgment, any representation of a fact (cf. *CPR*, B 131). Any representation of facts, such as “The train is green,” can be a conscious judgment: “*I think* that the train is green.” Otherwise we could have a representation without being able to recognize it as such. It does not necessarily need to be a conscious representation, but if it were somehow impossible to ascribe this representation to me, to become conscious of it at a later time, it would not count as a representation of a fact. Thus, the “I think” serves as a unifying ground for all my representations.

However, this “I think” is merely formal and does not refer to an actual being, but only to the formal subject of cognition. Kant’s transcendental deduction fails to deliver any insight into the reality of the subject inasmuch as a proper “I” still exists without being necessarily bound to thought. For Bouterwek, there is no doubt that the temporal existence of an actual mind or subject does persist even in absence of actual thinking, at least for a certain time (e.g. during moments of mere absent-mindedness or sleep). The question is: How can this “I” be captured conceptually in a way that goes beyond the merely formal “I” of the “I think”? This is where practical philosophy comes in, as it provides us with a concept of a subject that is independent of thought, for the cognitive act expressed in the “I will” provides us an indubitable, immediate experience of ourselves, of our way of reaching out to the world. Bouterwek writes:

If the thought: “I will” is to be more than [mere] imagination (*Einbildung*), then the I must certainly be thought of as something, namely as a living force: for if I will, then no other being wills but I. I am the source of all efficiency (*Wirkung*). I strive to submit nature to myself. (*Apodictic* II, 29)¹⁴

While on a formal level the “I think” and the “I will” seem similar, their referential object is very different. As Kant argues, experience is restrained to the realm of appearances – at least, Bouterwek would add, within the approach offered by theoretical philosophy. Nothing can be added to appearances to make them more real, since all additional (alleged) reality is restrained to the realm of thought. Let me try to illustrate this through an example: As Kant has argued, it makes no *theoretical* difference to imagine one hundred imaginative Thaler over imagining one hundred real Thaler. From the perspective of the cognizing subject, each is the same. But from the perspective of the willing subject, the difference is essential. I do not want to have one hundred imaginary Thaler; I want to have one hundred real Thaler. By willing, I immediately presuppose a meaningful distinction between me and someone else. *I* want to have one hundred Thaler. Not only do I want one hundred real Thaler, but I am also sure that it is *me* who wants to have them, not somebody else. Through willing, however, I have an immediate and non-propositional experience of both freedom and individuality that cannot be relegated to the realm of appearances.¹⁵ While I cannot explain *what* or *why*, I know *that* I am a free subject and, opposed to me, there is an objective reality. Whereas thought can be untethered from reality, the very act of willing presupposes the reality of both myself and the reality of my will. While a skeptic could argue that the formal subject of the transcendental apperception has no reality to it, the subject of the will cannot in the same way be a merely formal element of the subject-object relation. Instead, it necessarily refers to an actual object.

Obviously, we can have representations without willing the object of this representation to become reality. It is easy to conceive of a green train without willing the train to be green. Consequently, the will is not in the same way to be deduced a priori as thinking. The difference between the subject of the will and the transcendental “I” as the subject of cognition is that the will entails some activity that will inevitably be grounded in something, but not by force of the prop-

14 “Soll der Gedanke: ‘Ich will’ mehr als *Einbildung* sein, so muss das Ich durchaus als Etwas, und zwar als lebendige Kraft gedacht werden: denn wenn Ich will, so will kein anderes Wesen als Ich. Von mir geht alle Wirkung aus. Ich strebe, die Natur mir zu unterwerfen.”

15 Bouterwek occasionally identifies spontaneity with freedom (*Apodictic* II, 114) and, on other occasions, freedom is conceived as reason itself (e.g., *Apodictic* II, 108).

ositional content of the will, which does not have to be real. The will derives its objectivity not from the content of the will, but from its subject. Here, we can no longer conceive of a merely formal “I,” but must rather include an “I” that is real and alive. If the skeptic were to counter my claim to will something by casting doubt onto my existence, such an attack would be ultimately meaningless since, in the very act of willing, I am immediately aware of my existence – and I am such in a non-propositional way through an immediate experience of the *force* of the will.

While there are some parallels between Jacobi’s and Fichte’s approach to the will to be found, what makes Bouterwek’s approach so interesting and original is the idea that through the act of willing, reality becomes immediately and in a non-propositional way apparent as resistance to the will. In the following, I will sketch this argument, which is developed in the third book of the *Apodictic*.

The will is primarily seen as a motivational force that cannot be captured in concepts, just like reality itself (cf. *Apodictic* II, 32f.; 52). The deontological “I shall” is dismissed as derivative. However, from the perspective of practical philosophy, as it is conceived by Bouterwek, the will is unexplainable, an incomprehensible force of nature. But such a force is not conceivable without some resistance, which, in its own nature, must be force as well (cf. *Apodictic* II, 56 ff.; 68 ff.). To the extent that it is reality itself that resists our will, we need to conceive of reality as force(s) as well – an idea that can already be found in Locke and Leibniz. The force of the soul is just the same as the force of nature; but it is clearly not determined by nature, but rather by some irreducible spontaneity through which I become aware of my own reality. As we find ourselves as independent from the objective world, the will also serves as a ground for individuality. This is the source of how we come to conceive of ourselves as proper subjects, as the will allows us to establish the subject-object relation in a meaningful way in the first place (cf. *Apodictic* II, 73). From here, reality can be established as the proper foundation (*Realgrund*) of all knowledge, since through the act of willing I can be sure of both myself as a subject and my relation to something independent of myself, i.e. absolute reality.

How should we understand this titular resistance of reality against the will? The will assumes multiple functions in Bouterwek. While it serves as the ultimate principle of experience, it thereby also serves as a principle of the spontaneity of the mind, for the will determines the actual contents of our thought and our imagination. This is where Bouterwek’s distinction between thinking and knowing plays a big role that I want to illustrate by means of an example: Although I *cannot* will myself to accept the idea of a flat earth as true, I in fact *can* will myself to both imagine a flat earth and to think about it as much as I like. Here no resistance is encountered; this merely imagined and conceptualized flat earth

can be everything I want it to be. Apparently oblivious to the idea that coherence of thought might be constitutive to what I find to be true, Bouterwek assumes that it is reality itself that impedes me from willing this conception of a flat earth to be true. The correspondence of thought with reality is what we experience through this kind of resistance. The will thereby becomes the principle of reality (*Realprinzip*) that transcendental apperception failed to be; and it thereby also becomes the grounding of all philosophy (*Elementarprinzip*).

Through theory the 'I' is determined formally, but only through practice it is also determined materially, i.e. as a force, through the above mentioned non-propositional experience that goes along with each and every act of willing. Only with the inclusion of practical philosophy as a theory of the will can philosophy uncover an unshakeable and thereby absolute foundation for itself. From the viewpoint of practical thought, the "I," however, is not yet sufficiently determined as the subject of knowledge and consequently both theory and practice depend on each other for completion.

Reality has been defined by theoretical philosophy as the object of knowledge and by practical philosophy as the resistance to the will. Such a reality is, taken in itself and abstracted from any subject, not accessible for philosophy. We can only find the ultimate grounding of knowledge through the will as the mutual determination of the subject and the object and the apparent resistance of reality to the force of the will; this comprehensive system, which deals with the interrelatedness of subject and object, is a form of philosophy that Bouterwek calls "virtualism" or a "system of virtuality" (*Virtualsystem*) (*Apodictic* II, 71). This is supposed to allow us to conceive and experience a *virtual* unity of subject and object as forces (or rather: as force and resisting force) or as form and content.¹⁶ It can be called 'virtual' insofar as it precedes all claims of reality and thus cannot be conceived as real itself; but it is apparent in all experiences as the resistance of reality against our will. In this fashion, reality becomes apparent as the sum of all living forces.¹⁷ These ideas are later developed into a philosophy of nature strongly inspired by Schelling's (see Bouterwek 1802/1804).

16 This is also Bouterwek's criticism of Fichte's '*Tathandlung*:' there is no primitive force without opposition. However, Bouterwek's 'absolute virtuality' bears some resemblance to Fichte's conception of the absolute "I," which likewise contains a unity of both subject and object (*Apodictic* II, 70).

17 This should not be misunderstood as some kind of vitalism: living forces can only be conceived as the appearances of absolute reality, which is nothing but an all-comprehensive unity. Forces should not be understood as independent entities (*Apodictic* II, 69: "Absolute Kräfte, Urkräfte, Grundkräfte *an sich*, oder wie man sie sonst nennen will, sind Undinge"). Instead, they are relative (see *Apodictic* II, 208).

The will does the heavy lifting in Bouterwek's philosophy. It is the principle of knowledge as well as the principle of thought (i.e. judgment) (cf. *Apodictic* II, 105) inasmuch as it provides a criterion to distinguish between both and it also serves as a causal ground for the production of any actual thought. By enabling us to distinguish between subjects and objects, it is also a principle of subjectivity (cf. *Apodictic* II, 119; 127). Beyond that, it also serves as a principle of freedom given that the will allows us to establish ourselves as independent from objects and their connections: We are free because we can *think* independently of the objects; we 'break free' from our objects (*von den Objekten losreißen*) whereas our knowledge is bound to reality. Thereby we find ourselves as independent.¹⁸ This is not the theoretical freedom that Kant calls transcendental freedom, in which we necessarily have to conceive of ourselves as free so that we can deliberate about moral issues in a meaningful way; this is a practical freedom that needs to account for reality as it is experienced as an opposing force.¹⁹ Reason plays a role in our agency insofar as it allows us to make a choice between opposing impulses; only then can we realize the freedom of the will by means of reason (see *Apodictic* II, 132ff.). Unless we account for the practical limitations of our actions in the world, all conceptions of 'pure spontaneity' or '*reine Tathandlung*' are meaningless. Only the will provides us with certainty about the existence of an objective world. It does not provide any content to our perceptions, as that is the task of imagination, which transforms virtuality (the virtual unity of subject and object, of will and resistance) into perceptual content by subtracting the subjective side from the mutual determination. Theoretical and practical philosophy are thus supposed to be unified on grounds of the non-propositional experience of the will, which also serves as a ground for the subject-object relation, i.e. our conception that appearances are not merely subjective imaginations, but rather appearances *of something*. This something, however, still needs to be determined through knowledge and thought, wherein we distinguish between subject and object.

Similarly, the experience of our freedom of being able to 'tear ourselves away from objects' is supposed to lead us into a conception of duty – in short, the alleged foundation of practical philosophy is related to ethical claims. This is because, for Bouterwek, the appearance of reality is intermixed with a presumptive "spirit world" (*Geisterwelt*, cf. *Apodictic* I, 311; II, 160) in which ethical delibera-

18 "Wir sind vernünftig dadurch, dass wir uns von den Objekten losreißen und uns selbständig finden. Wir sind frei dadurch." (*Apodictic* II, 113)

19 Bouterwek's use of the term 'practical' may be irritating here. Perhaps it could best be understood as the agent's freedom or freedom to act in accordance to the will. Apparently, it does not yet entail any normativity.

tions take place. This is not so much a result of rational deliberation, but rather given as a feeling in practical approaches to everyday situation. Because we all experience our own will, it is fair to assume that other beings that resemble us will also experience a similar will and therefore constitute their own individuality. While we cannot prove such an assumption from experience, we are nevertheless enticed to assume a plurality of subjects. Here, Bouterwek shoehorns some foundational remarks on ethics into his transcendental philosophy that involves a rather underdeveloped understanding of duty, but here is not the place to discuss these consequences further.

In the end, Bouterwek claims to have established the unity of theoretical and practical philosophy on the foundation of the non-propositional elements of our experience of reality through the will. In doing so, he has rescued the claims for objectivity from the attacks of the skeptic. Admittedly, this is more of a catchphrase than a meaningful philosophical statement; and Bouterwek does not follow his own conception of the foundations of practical philosophy through into a system of ethics. The only conclusion he draws that bears some resemblance to ethical deliberations is rather trivial: By ‘saving’ reality from skeptical attacks, Bouterwek redefines the task of theoretical philosophy as “Find yourself as thinking in the world.” (*Apodictic* II, 227)²⁰.

6 Conclusion

After this brief overview of Bouterwek’s ambitious, but often divergent and convoluted book on transcendental philosophy, which spans more than 700 pages, I would like to conclude with some general remarks on the advantages and disadvantages of his project.

Let me first examine some of the explanatory advantages that Bouterwek has to offer. First, his criticism of Kant and the transcendental project, especially the failure to distinguish between knowledge and thought and the corresponding failure to relate thought to existence is, as even Erich Adickes concedes in his otherwise extremely critical assessment of the *Apodictic*, “not without merit.”²¹

As Bouterwek aims to secure the core ideas of Kant’s first *Critique* against possible attacks on the part of the skeptic, he chooses a different starting-point. Kant’s approach consists in understanding the conditions that make metaphysical knowledge possible. Kant argues that, insofar as mathematics is obvi-

²⁰ “Finde denkend dich selbst in der Welt.”

²¹ Adickes 1895, 409.

ously possible (even a skeptic like Hume had not doubted the validity of mathematics) and metaphysics follows the same structure, i.e. consists in synthetic a priori judgments, we can reasonably assume that metaphysics is equally possible. For Bouterwek, such a line of reasoning is still vulnerable to other skeptical attacks; it would, for example, not dissuade a Cartesian skeptic from rejecting the validity of both mathematics and metaphysics just alike. Consequently, we need to abandon the Kantian presupposition that sets his critical enterprise in motion: *Assuming that we know something...* Bouterwek instead suggests falling back on a slightly different starting-point: *Assuming that something exists...* This ‘something,’ of course, is apparent as something outside of me and it reveals itself in the rather trivial fact that it resists our will: we don’t always get what we want, nor can we choose what to believe. Experience, even the non-propositional contents of our experience, can provide us with an indication that other things besides me exist and that they lie beyond my control as a rational subject. Whether the radical Cartesian skeptic would find such an approach entirely plausible is doubtful, but nonetheless Bouterwek has provided a counterproposal that is entirely new.

His counterproposal also has the advantage of strengthening the relevance of the non-conceptual content of human experience against what we may call the “concept fixation” of transcendental epistemology, i.e. the idea that epistemology can defend itself by means of some propositional content against all possible skeptical attacks. For Kant, intuitions cannot guarantee any objectivity, only the categories can. Against this Bouterwek insists that rational judgments are not made in an otherwise ‘empty’ space of reasons, but that they, depending on their scope and content, exist within a comprehensive framework that needs to contain non-propositional aspects of our experience – a line of thought that is not alien to other modern and contemporary philosophers.

Bouterwek also allows us to pick up some core ideas of moral sentimentalism and integrate them into a broader framework of transcendental philosophy. It is not adequate to reduce the human being to a mere rational subject as we should not ignore the impact of affections on the will. In this regard, while Bouterwek agrees with Kant that such causal determinations can never serve as valid ethical justifications for our actions, he insists that the mere fact that such a causal external determination exists requires of us that we reconcile our given human nature with the claims of deontic ethics. Here, Bouterwek anticipates Heidegger’s account of our constitutive “being-in-the-world,” namely the latter’s theory of the subject as that which is never merely a pure subject of mental states, but rather an embodied being in a complicated network of relations to other beings.

Finally, Bouterwek offers us a conception of the will as a natural force that can be conceived as sufficiently similar, ontologically speaking, with other natural forces. This allows us to establish a strong relation between the transcendental freedom of the will and the freedom to act as an embodied being – something that Schopenhauer will pick up on later and which will also be repeated by other French and German philosophers of the 19th century, including Marx, Dilthey and Bloch.²²

Now, let's take a look at the problems that arise. It is rather unconvincing to assume that the experience of the will and the resistance of the world to the will could serve as a criterion for truth. Bouterwek strives to make a non-conceptual feature of our cognition to the ground of all conceptual knowledge – in a sense, it serves as a 'tether' to reality. But he does not offer any clues as to how this bridge between the conceptual and the non-conceptual could be conceived within an explanation or justification. The mere fact that *something* must be *out there*, even if this fact is conveyed through immediate experience, does not translate, without further argument, to a foundation of propositional knowledge. The resistance of reality against the will cannot be communicated and is therefore useless in actual justifications. Even if I am willing to concede the point that reality can be *felt* as a resistance, there is little that could I answer in response to the Cartesian skeptic beyond "Don't you feel it too?". The Cartesian skeptic could easily denounce this as metaphysical mysticism, a vague belief in an essentially unknowable reality that could still be attacked with the thought experiment of the malign spirit that induces our perceptions and feelings from the outside, without allowing for any guesses about what the respective cause of these perceptions and feelings might look like.

Bouterwek also employs an identity relation between mental spontaneity and physical forces that sounds profoundly Leibnizian, and it is not easy to see how such an idea could be developed within the framework of post-Kantian thought. He does not allow for an individuation of forces, although he insists that the will lets us experience our individuality. Similarly, this system of virtuality does not consist of a plurality of subjects and objects: we can experience ourselves only as a generic subject through the will (i.e. a subject in general) and consequently the experience of a resisting reality allows for nothing but a postulate of an object in general. Getting from these general concepts to the assumption of an externality of differentiated and individual subjects and objects re-

²² See Schmidt 1985. Schmidt emphasizes that Bouterwek was the first one to develop a version of this argument that only later became more popular. Bouterwek did, however, not yet fully realize that such an oppositional relation to reality can have an orienting and structural function to our self-conception as humans – this is what other, later authors emphasize instead.

quires a leap of faith that is grounded only by some rather incoherently inserted remarks about God. It thereby shifts the conception of a subject (or object) in general to a conception of particular subjects (or objects) that is not warranted. In short, his ontology is in this respect rather muddled and the relation between mental spontaneity and physical forces is not at all clear.

But perhaps the greatest disadvantage of Bouterwek project is a certain ambiguity it led to. By emphasizing the importance of skepticism and the profound failure of his contemporaries to provide some secure grounds against it, Bouterwek was, in fact, seen as someone who had rather embraced skepticism himself. He does not refute this interpretation, but embraces it. He writes:

Do you know what Villers²³ calls our philosophy? *Sentimental skepticism*. I like that title. Because “sentimental” in its original meaning is what comes from the heart; and skepticism is a teaching that belongs to pure understanding. Sentimental skepticism would thus be a skeptical doctrine against which the heart, otherwise called practical reason, can’t argue, but which it rather approves. (Letter to Jacobi, 24. Feb. 1801. In: Jaeschke 1999, 145)^{24,25}

Bibliography and Abbreviations

- Apodictic I*: Friedrich Bouterwek: Idee einer Apodiktik, Teil I (1799). Ed. by Ansgar Lyssy. Stuttgart: Frommann-Holzboog, 2018.
- Apodictic II*: Friedrich Bouterwek: Idee einer Apodiktik, Teil II. Halle: In der Rengerschen Buchhandlung, 1799.
- AA: Immanuel Kant: *Gesammelte Schriften*. Ed. by the Prussian Academy of Science. Berlin/New York: De Gruyter, 1900-.
- Adickes, Erich (1896): *German Kantian Bibliography*. Vol. 3. Boston: Athenaeum Press, pp. 407–414.
- Blumenbach, Johann Friedrich (1828): “Memoria Friderici Bouterwek,” in: *Commentationes Societatis Regiae Scientiarum Gottingensis Recentiores* 7, pp. 3–10.

23 Charles François Dominique de Villers (1765–1815), a French philosopher famous for translating Kant into French. He later in 1806 became Bouterwek’s colleague at the university of Göttingen.

24 “Wissen Sie, wie Villers unsere Philosophie betitelt? *Sentimentalen Skeptizismus*. Der Titel gefällt mir. Denn sentimental in der ursprünglichen Bedeutung ist, was von Herzen kommt; und der Skeptizismus ist eine Lehre, die dem puren Verstande angehört. Sentimentaler Skeptizismus wäre also eine Zweifelslehre, gegen die das Herz, sonst genannt praktische Vernunft, nichts zu erinnern hat, die es vielmehr sanktioniert.”

25 I would like to thank Joseph Carew for his helpful comments and corrections.

- Bouterwek, Friedrich (1801–1819): *Geschichte der Poesie und der Beredsamkeit seit dem Ende des dreizehnten Jahrhunderts*. 12 volumes. Göttingen: Johann Röwer.
- Bouterwek, Friedrich (1802/1804): *Von der Naturphilosophie nach dem System der Apodiktik*. In: *Neues Museum der Philosophie und Literatur*. Vol. I.1 (1802) and Vol. II.2 (1804).
- Bouterwek, Friedrich (2018): *Idee einer Apodiktik, Teil I* (1799). Ed. by Ansgar Lyssy. Stuttgart: Frommann-Holzboog.
- Bouterwek, Friedrich (1799): *Idee einer Apodiktik, Teil II* (1799). Halle: In der Rengerschen Buchhandlung.
- Dierse, Ulrich (1999): "Bouterweks Idee einer Apodiktik." In: Jaeschke, Walter (Ed.): *Der Streit um die Gestalt einer Ersten Philosophie (1799–1807)*. Vol. II. Hamburg: Meiner, pp. 32–51.
- Henrich, Dieter (1991): *Konstellationen: Probleme und Debatten am Ursprung der idealistischen Philosophie (1789–1795)*. Stuttgart: Klett-Cotta.
- Jaeschke, Walter (1999) (Ed.): *Der Streit um die Göttlichen Dinge (1799–1812)*. Vol. III. Hamburg: Meiner.
- Kant, Immanuel (1900-): *Gesammelte Schriften*. Ed. by the Prussian Academy of Science. Berlin/New York: De Gruyter.
- Ludwig, Bernhard (1988): *Kants Rechtslehre*, Hamburg: Meiner.
- Lyssy, Ansgar (2014): "Kant für Jedermann. Über F. Bouterweks Versuch, Kants kritische Philosophie populär darzustellen." In: Christoph Binkelman / Nele Schneidereit (Eds.): *Denken fürs Volk – Formen von Gemeinsinn in der deutschen Popularphilosophie*. Würzburg: Königshausen & Neumann, pp. 139–167.
- Melo Araújo, André de (2012): *Weltgeschichte in Göttingen: Eine Studie über das spätaufklärerische universalhistorische Denken 1756–1815*. Bielefeld: Transcript.
- Pfaff, Carl (1853) (ed.): *Sammlung von Briefen gewechselt zwischen Johann Friedrich Pfaff und Herzog Carl von Württemberg, F. Bouterwek, A. v. Humboldt, A.G. Kästner, und Anderen*. Leipzig: J. C. Hinrichs Buchhandlung, 1853.
- Schmidt, Burkhard (1985): *Das Widerstandsargument in der Erkenntnistheorie*. Frankfurt a. M.: Suhrkamp.
- Senne, Linda Lou Prusiecki (1972): *Friedrich Bouterwek, the philosophical critic. An intellectual biography*. PhD Thesis, Microfiche. Stanford University.
- Struck, Gustav (1919): *Friedrich Bouterwek, sein Leben, seine Schriften und seine philosophischen Lehren*. Rostock: Carl Hinstorffs Hofbuchdruckerei.
- Zinkstok, Job (2010): Art. "Bouterwek." In: Heiner Klemme et al. (eds.): *The Dictionary of Eighteenth-Century German Philosophers*. Vol. 1. London: Bloomsbury, pp. 137–141.
- Zymner, Rüdiger/ Hölter, Achim (2013) (eds.): *Handbuch Komparatistik*. Stuttgart: Metzler.

Jörg Noller

“Will is Primal Being”: Schelling’s Critical Voluntarism

Abstract: The aim of this paper is to analyze Schelling’s conception of will particularly in his *Philosophical Investigations into the Essence of Human Freedom* (1809). Whereas many interpreters tend to understand his theory as an obscure metaphysics of will, I shall argue that Schelling does not hold a metaphysical voluntarism, but rather develops what I shall call a “critical voluntarism.” According to Schelling’s critical voluntarism, the will is deeply connected to the faculty of reason although it is not identical with it as Kant seems to suggest. Schelling transforms Kant’s conception of autonomy into what he calls a “real and vital concept” of freedom that is, according to him, “the capacity for good and evil” (FE, 352). Schelling’s critical voluntarism entails what I shall call a “real compatibilism,” according to which being a natural entity does not rule out our freedom of the will but rather grounds it. I shall finally analyze Schelling’s conception of freedom by referring to Harry Frankfurt’s theory of volitional necessity.

1 Introduction

The aim of this paper is to analyze Schelling’s conception of will particularly in his *Philosophical Investigations into the Essence of Human Freedom* (1809). Whereas many interpreters tend to understand his theory as an obscure metaphysics of will—or, as Heidegger did—a “metaphysics of evil” (GA 42, 168)—I shall argue that Schelling does not hold a metaphysical voluntarism as Schopenhauer did in his *World as Will and Representation*, but rather develops what I shall call a “critical voluntarism” by dealing with Kant’s conception of freedom.¹ According to Schelling’s critical voluntarism, the will is deeply connected to the faculty of reason although it is not identical with it as Kant seems to suggest.² Schelling transforms Kant’s conception of autonomy into what he calls a “real

¹ Recently, the Kantian roots of Schelling’s theory of freedom have been emphasized. See Gardner 2017, 134: “Schelling’s central claims in the *Freiheitsschrift* can be regarded as the product of a complex and extended development arising out of Kant’s theory of freedom.”

² According to Kant’s *Groundwork of the Metaphysics of Morals*, the will is “nothing else than practical reason” (GMM, 4:412).

and vital concept” of freedom that is, according to him, “the capacity for good and evil” (FE, 352).

The paper has three parts. First I will outline Kant’s theory of autonomy, referring to the tight relationship between will and practical reason. I will then outline the so-called “Reinhold’s dilemma” (Allison 1986, 422) or “Reinhold’s complaint” (Kosch 2006, 55) that concerns our freedom to act immorally if the will is identified with practical reason. After Kant, thinkers such as Carl Christian Erhard Schmid took this identification seriously and argued for an intelligible fatalism, according to which only morally good actions are free, whereas morally evil actions are determined by external influences of sensibility. Finally, I shall situate Schelling’s philosophy within the debate on freedom after Kant.³ I will therefore consider Schelling’s attempt to solve Reinhold’s dilemma in his *General Overview of the Most Recent Philosophical Literature* (GO) from 1798, in his *System of Transcendental Idealism* (STI) from 1800, and his later *Freedom Essay* (FE) from 1809.

2 Kant’s challenge

Kant’s theory of human freedom is motivated by the strong claim of a moral agent’s absolute volitional imputability. He claims that “the human being must make or have made himself into whatever he is or should become in a moral sense, good or evil. These two [characters] must be an effect of his free power of choice (*Willkür*), for otherwise they could not be imputed to him and, consequently, he could be neither morally good nor evil” (RBMR, 6:44). In order to argue for his imputability claim, Kant needs to do justice to two requirements.⁴ On the one hand, a person’s decision must not depend on *external* factors, for in this case an action would not be imputable to the person’s will (i.e. heteronomy). On the other hand, a free decision must not be groundless (i.e. indifferentism), but must follow from reasons that stem from the inner person’s will. Kant attempted to fulfil both requirements by introducing his concept of a “pure will” (RBMR, 6:45). The pure will is essentially independent from external influences, and depends only on its *own* laws (i.e. autonomy).

In order to guarantee absolute and radical self-determination, Kant refers to his conception of freedom as autonomy. In his *Critique of Practical Reason*, which

³ Recently, there has been growing interest in this debate. See Fugate 2015, Gardner 2017, Kosch 2006, Martin 2017, Noller 2015, Noller 2019a, and Ware 2019.

⁴ See Noller 2015, 16–23.

—even more than his *Groundwork*—is concerned with this problem, Kant formulated the “first question” in the sense of “whether pure reason of itself alone suffices to determine the will or whether it can be a determining ground of the will only as empirically conditioned.” (CPrR, 5:15) This question arises from the critical position of the human will, which “stands between its a priori principle, which is formal, and its a posteriori incentive, which is material, as at a crossroads” (GMM, 4:400). Absolute freedom of the will is only possibly by the lawful form of pure, that is empirically non-contaminated reason. To decide on the basis of material, that is concrete and contingent motives, however, would obliterate its autonomy insofar as “all laws that are determined with reference to an object give heteronomy” (GMM, 4:458). The first requirement of autonomy is therefore the *independence* of any material motives, hence “the first concept of it is negative” (CPrR, 5:29).

The crucial point of autonomy, however, is its positive freedom—not freedom *from* but *to*. Negative freedom alone does not suffice to explain the full concept of an autonomy of the will: “The preceding definition of freedom is *negative* and therefore unfruitful for insight into its essence; but there flows from it a *positive* concept of freedom, which is so much the richer and more fruitful.” (GMM, 4:446) An entirely indifferent and unlawful will would be independent from the law of nature, however, the will’s decision would not have any determination at all and would be arbitrary in the bad sense (i.e. indifferentism). For that reason, Kant insists that freedom of the will “is not for that reason lawless but must instead be a causality in accordance with immutable laws but *of a special kind*; for otherwise a free will would be an absurdity” (GMM, 4:446; my emphasis).

Kant explains this special kind of law in terms of a special kind of *causality* that he refers to as a “causality of reason” (CPrR, 5:80) or “causality through freedom” (CPrR, 5:47) (see Noller 2019b). It is the will under the moral law that establishes such a causality as “a true *higher* faculty of desire, to which the pathologically determinable is subordinate” (CPrR, 5:25). According to Kant, the “[w]ill is a kind of causality of living beings insofar as they are rational, and freedom would be that property of such causality that it can be efficient independently of alien causes *determining* it” (GMM, 4:446). This entails the “reciprocity thesis” (Allison 1996), according to which “a free will and a will under moral laws are one and the same” (GMM, 4:447).

3 Reinhold's dilemma

Kant's theory of autonomy raises a serious issue when it comes to moral imputability.⁵ This "imputability problem" (Hudson 1991, 179) stems from a conflict between Kant's general imputability thesis (IT) and his autonomy thesis (AT), which shall justify (IT).

(IT): The free agent is morally responsible for her morally right and wrong actions and has free choice between the alternatives of good and evil.

(AT): The absolute cause of the autonomous action lies in the causality of pure practical reason and its moral law.

From IT and AT follows the so-called "autonomy problem" (AP), which can be explicated in a strong (AP1) and a weak sense (AP2).

(AP1) A causality of free action that contradicts the moral law cannot be *thought* consistently for this causality itself stems from pure reason. The category of evil cannot be consistently explained as a "modus" (CPrR, 5:65) of a causality of freedom on the "basis" (Kant, RBMR, 5:66) of the moral law; hence, an evil action cannot be a product of autonomous reason, and therefore *is not* an autonomous action.

(AP2) A practical cognition of evil means to decide *at the same time* to give up one's autonomy by jumping from autonomy to heteronomy. This jump itself, however, cannot be explained within the framework of a conception of autonomy and therefore seems to be a groundless *event*. A voluntary free jump to heteronomy falls short of autonomy, for it happens within a lawless sphere.

In Anglophone scholarship, Henry Sidgwick was the first to hint at AP1: "Kant, either expressly or by implication, identifies Will and Reason; for this identification obviously excludes the possibility of Will's choosing between Reason and non-rational impulses." (Sidgwick 1888, 411) Sidgwick identified a "confusion" between two different conceptions of freedom in Kant's theory, namely "(1) the Freedom that is only realised in right conduct, when reason successfully resists the seductions of appetite or passion, and (2) the Freedom to choose between right and wrong, which is, of course, equally realised in either choice." (Sidgwick 1888, 405)

Almost one hundred years earlier, however, Karl Leonhard Reinhold had already pointed to AP1 in the Second Volume of his *Letters on the Kantian Philos-*

⁵ See Noller 2019a, 854–855.

ophy from 1792. Especially in the Sixth, Seventh and Eighth Letter, Reinhold provides a subtle analysis of the imputability problem, and then attempts to solve it by introducing a modified action theory as well as an elaborated concept of freedom of the will in order to perfect Kant’s account of autonomy. Already in his letter to Jens Immanuel Baggesen, dating from March 28 of 1792—immediately before the First Section of Kant’s *Religion* appeared—he gave an overview of his most important points of criticism, and also revealed his new and different understanding of freedom as *individual* autonomy in opposition to Kant’s general account:

I utterly distance myself from Kant and the Kantians concerning the concept of will, which I neither take for a causality of reason, nor a faculty to act according to given laws and so forth, but for a person’s faculty, equally distinct from reason and sensibility, to determine oneself towards the satisfaction or non-satisfaction of a desire (claim of the self-interested drive). (Reinhold 1831, 168)

After this mostly negative definition, Reinhold goes on specifying his concept of *positive* freedom:

Indeed, this self-determination takes place by rules, and insofar by reason; *but thereby reason behaves as a mere faculty that can be used by the willing subject in two different ways*, for it depends on the subject whether the rule is used as a mere means for the satisfaction of the selfish drive, or as an end [...]. (Reinhold 1831, 168f.)

Aside from the modified concept of the use of reason, we find a crucial modification of the relationship between empirical and rational motives of the will here. For Reinhold adds in the aforementioned letter: “I even distance myself from Kant concerning the concept of morality [...] for I cannot conceive of morality without *sensibility*.” (*Baggesen Letters 1*, 168)

Reinhold also highlights some problematic implications that seem to follow from Kant’s theory of the autonomy of reason. The “friends of Kantian Philosophy,” as Reinhold puts it, “attempted to save the will from the slavery of the instinct only in that way insofar as they made it the slave of the force of reason.” These “friends,” according to Reinhold, “attempted to escape will’s necessitation of sensibility only by conceiving the will as being inevitably necessitated by reason” (Reinhold, *Letters II*, 200). According to this conception, however, “a moral action could only be understood as a mere effect of reason” (Reinhold, *Letters II*, 200). If “the will were only free with regard to *moral* actions, and the ground of *immoral* actions laid outside the will in external obstacles and barriers,” Reinhold goes on,

so also the reason of *moral* actions would by no means to be found in the mere self-activity of practical reason but rather in the *absence* of these obstacles that are entirely independent from this reason. The whole freedom of this reason, and by it also of the person, would only consist in an accidental independence from external force and restricted to certain cases that would by no means be in the control of the person. The moral action would follow inevitably by a entirely involuntary effect of practical reason *as soon as there would not be any obstacle*; and the moral or immoral action would have to be imputed to the sheer presence or absence of this obstacle. (Reinhold, *Letters II*, 200)

Intelligible fatalism and empirical determinism are, according to Reinhold, only two sides of the same coin for neither can explain how a person can both rationally and freely decide to act against the moral law. Hence, the conceptual challenge for Reinhold consists in developing a complex concept of individual freedom that avoids (i) intelligible fatalism, (ii) empirical determinism and (iii) indifferentism. The faculty of the will must, on the one hand, not be hypostatized and conceived as something absolutely distinct and separated from reason, which would lead to indifferentism. On the other hand, it must not be fully identified with reason, which would lead to a rational determinism that would render any voluntary deviation from the moral law conceptually impossible.

4 Kant's reply

The importance of Kant's late *Metaphysics of Morals* consists in the fact that it provides his final statement about Reinhold's dilemma, attempting to clarify the relation between will, power of choice, reason and moral law. For this reason the concept of the power of choice (*Willkür*) is of special interest. Kant attempts to situate it within his system of transcendental philosophy and to clarify its determining ground. This is a direct reaction to Reinhold's insisting on the possibility of *positive* freedom *against* the moral law. Kant's final remarks on the autonomy problem can be found in the fourth section of the *Metaphysics of morals*; they encompass roughly one single page (MM 6:226 f). They read more like explanations or annotations than a genuine theory. The fact that these final remarks are very short, however, must not hide the fact that they are of special systematic importance concerning Reinhold's dilemma (see Noller 2019a, 858–861).

Like in his *Religion*, Kant defines individual freedom as a maxim, namely as the “rule that the agent himself makes his principle on subjective grounds” (MM 6:225). Moreover, Kant defines the concept of a deed as an action “insofar as it comes under obligatory laws and hence insofar as the subject, in doing it, is considered in terms of the freedom of his choice (“Willkür”)” (MM 6:223). It is striking that Kant defines the “capacity for desiring in accordance with concepts” as

the “*capacity for choice*” (“Willkür”) (MM 6:213), and not, as he did in his *Second Critique*, as the “higher volitional capacity” (“höheres Begehrungsvermögen”), that is, pure practical reason (CPrR 5:24f.). At this point, Kant sharply distinguishes between will (“Wille”) and faculty of choice (“Willkür”):

The capacity for desire whose inner determining ground, hence even what pleases it, lies within the subject’s reason is called the will. The will is therefore the capacity for desire considered not so much in relation to action (as the capacity for choice is) but rather in relation to the ground determining choice to action. The will itself, strictly speaking, has no determining ground; insofar as it can determine the capacity for choice, it is instead practical reason itself (MM 6:223).

It follows from this that it is not pure practical reason that immediately produces an action as the *principium executionis* through the moral feeling of respect; rather, the faculty of the power to choose takes up this intermediate position between will and action. Hence, Kant situates the faculty of choice *below* the will as a kind of faculty to realize the will in the empirical world (cf. MM 6:213).

This definition of the faculty of choice results in the will only having a legislative function, whereas the power to execute the will lies in the faculty of choice: “Laws proceed from the will, *maxims* from choice (“Willkür”).” (MM 6:226) Since the will “is directed to nothing beyond the law itself, [it] cannot be called either free or unfree, since it is not directed to actions but immediately to giving laws for the maxims of actions (and is, therefore, practical reason itself). Hence the will directs with absolute necessity and is itself subject to no necessitation.” (MM 6:226) It follows that “[o]nly choice (“Willkür”) can [...] be called free.” (RBMR 6:226) This is due to the fact that it can choose the concrete *matter* of the maxim, whereas the will *qua* moral law always provides the *form* to which the faculty of choice can freely refer—a distinction that is likely due to Reinhold’s dilemma.

But how can we make sense of Kant’s statement that the will “cannot be called either free or unfree”? We find further arguments in Kant’s preliminary version (“Vorarbeiten”) to the preface and introduction to the *Metaphysics of Morals* (23:243–252, here 248–250) where he insists on the separation between will (“Wille”) and choice (“Willkür”).⁶ The will cannot be called free since it is “not under the law but [...] itself the legislator for the faculty of choice and [...] absolute practical spontaneity determining the faculty of choice.” (23:248) For this reason, the will is “good in all human beings, and there is no unlawful willing” (23:248). The faculty of choice stands under the will, which is the instance of leg-

⁶ “Man’s will must be distinguished from the faculty of choice (“Willkür”)”; Kant, 23:248.

isolation. Insofar as it is not conceptually bound to the moral law, Kant conceives of it as a “natural faculty” whose maxims “since they can refer to actions as appearances in the empirical world can be evil.” As Kant puts it, the faculty of choice is “actually not immediately determinable,” “but only by means of the maxims to use it according or against the moral law” (23:248).

A closer look, however, reveals that Kant attributes a specific form of freedom to the will. Whereas the faculty of choice, as a phenomenon, is free to materially decide for or against the moral law, “the will [...] is free in another sense since it is lawgiving not obeying neither the law of nature nor another one, and insofar freedom is a positive faculty, however not to choose, since there is no choice, but to determine the subject with regard to the empirical action.” (23:249) Given the ‘critical’ position of the faculty of choice within Kant’s transcendental idealism, his thesis that “[o]nly freedom in relation to the internal lawgiving of reason is really a capacity; the possibility of deviating from it is an incapacity” becomes clear (Kant, MM 6:227). But the fact that Kant speaks of “incapacity” is not equivalent to “impossibility,” as some interpreters have held. Rather, impossibility means in this context a *privative* form of the faculty of freedom as the freedom to act immorally.

Insofar as there is no possibility to *explain* the subjective law of the capacity of choice from the perspective of practical reason, there remains only the possibility to define the morally good, which is extensionally identical to an action out of autonomous reason. This means that Kant’s final conception of autonomy does not entail the strong AP1 but the weaker AP2. But AP2 does not entail an intelligible fatalism, which would only follow if one understood the epistemic and definitional restriction of the evil capacity of choice *metaphysically*.⁷ Rather, Kant’s epistemological escape from intelligible fatalism consists in the conceptual distinction between *incapacity* and *impossibility* of evil actions, the former guaranteeing their *logical* possibility.

Even if Kant avoids intelligible fatalism, his theory of individual freedom remains highly problematic, since a morally evil action seems to be *less free* than a good one, and, hence, *less imputable*. In fact, Kant can hardly explain how—on the basis of the moral law—a free agent should be possible to produce an evil action as a sort of practical cognition, as he argued in his Second Critique’s chapter “On the concept of an object of pure practical reason” (RBMR 6:57–71). Against such an asymmetric account the conceptual challenge consists

7 Fugate seems to read Kant in this way: “He [Kant] uses rather the word ‘Unvermögen’ which most certainly means an absence of a power or faculty, or more precisely, an impossibility of acting in a specifically defined way” (Fugate 2015, 362).

in developing a symmetric concept of freedom according to which good and evil actions are likewise and to the same degree free and imputable. Moreover, the conceptual challenge consists in developing an account of individual freedom that is neither an indifferent occurrence (“indifferentism”) nor a necessarily determined event (“intelligible fatalism” / “empirical determinism”).

5 Freedom “from the standpoint of consciousness”: Schelling’s early dispute with Kant and Reinhold

In his *Metaphysics of Morals*, Kant argues that freedom cannot be *defined* as acting against the moral law, and he also argues that acting against the moral law is due to an incapacity (“Unvermögen”). Schelling, however, attempts to determine freedom against the moral law in a positive way and to conceive of freedom to evil as a *capacity*.⁸ Already before his *Freedom Essay* was published in 1809, Schelling had discussed Reinhold’s dilemma and Kant’s reply in his *General Overview of the Most Recent Philosophical Literature* (*Allgemeine Übersicht der neuesten philosophischen Litteratur*) from 1798, and further developed it in his *System of Transcendental Idealism* from 1800. Schelling’s theory aims to “balance the antagonism that seems to take place in the propositions of two famous philosophers on this subject.” (GO 157) This antagonism consists in the opposed conceptions of freedom as autonomy and as choice (“Willkür”) given by Kant and Reinhold respectively (see Noller 2015, 300–309).

How does Schelling’s early philosophy relate to both theories of freedom of the will? Schelling’s conception can be interpreted as an attempt to defend Kant’s account against Reinhold’s charge, but at the same time to reconstruct Kant’s problematic statements about the concept of will in such a way that it can withstand Reinhold’s critique. Reinhold was, as Schelling puts it, “not determined to *solve* the very problem of philosophy but to *represent* it in the most determined way.” (I, 99)

The Kantian standpoint of the absolute will and pure practical reason is, according to Schelling, superior to Reinhold’s, since we can derive the concept of choice (“Willkür”) from it. Reinhold is not able to analyze the concept of will on the basis of his conception of choice (“Willkür”), since he can only “appeal to

⁸ See Gardner 2017, 142: “Evil is [...] for Kant a negative, indirect condition of human freedom which does not belong directly to its very concept – as it does for Schelling.”

the judgement of common practical reason, which he cannot explain any further.” (GO, 162) Whereas Reinhold’s analysis remains in the sphere of empirical facts, Kant, as Schelling puts it, can “proof from principles” that the will must “appear as free choice (“Willkür”),” “even though this capacity is not conceivable at all in the absolute will (that is lawgiving)” – which is the reason why Kant had called the choice (“Willkür”) of evil an “incapacity.” (Kant, MM 6:227), Schelling therefore criticizes Reinhold explicitly not for conceiving real freedom as freedom of choice but for the philosophical standpoint from which he attempts to analyze it. For Reinhold derives his conception of freedom from the “ordinary consciousness” (“gemeines Bewusstsein”), that is, he does not situate the concept of freedom sufficiently within a philosophical system, which, according to Schelling’s early philosophy, must always begin with the absolute.

With the concept of choice (“Willkür”), the conception of finite and individual freedom becomes more and more important to Schelling’s philosophy. Indeed, in his treatise *Of the I as Principle of Philosophy* (I) from 1795, Schelling already discussed finite freedom. He had, however, qualified the “final aim of all striving” to be “the expansion of personhood towards infinity” and argued for its “destruction” (“Zerstörung”) for this purpose (I, 128). The intent of Schelling’s conception of freedom in the wake of Kant and Reinhold is to develop a conception of individually imputable will from the standpoint of the absolute will, that is “to make the consciousness of freedom comprehensible (as it were to construct it),” which shall be realized “by the concept of choice” (GO, 163).

Schelling does not, however, identify freedom with the absolute will – as he did afterwards – for this would imply the danger of intelligible fatalism, but relates both entities:

For the explanation of free choice (as a fact of common consciousness) we need the idea of absolute freedom; without it we cannot comprehend any freedom of choice; with it alone we do not understand how a choice still be possible at all in us, and why the original law has not become necessity in us. (GO, 165)

The phenomenon of choice is, according to Schelling, the *ratio cognoscendi* of freedom, whereas Kant had related choice to the moral law itself: “That there is a freedom of will, thereof the common consciousness is convinced by the capacity of choice, that is by the fact that in each willing we become aware of a choice between oppositions.” (STI, 275) According to Schelling, it follows from this that “freedom = choice” (STI, 275). The absolute will, as a mere determinability, must be determined to get to a decision. This is, however, only possible in the form of choice as the appearance of the absolute will: “Choice,” as Schelling

puts it, "is necessary for the possibility of the representation of our free action" (GO, 157 n.).

How does Schelling solve the dispute between Kant and Reinhold on freedom of the will? As a first step, he opposes the two concepts of freedom in order to find the reason for their dissent: "*Kant* argues in his Critique of practical reason that will and practical, that is law-giving reason be one and the same." And Schelling goes on: "*Reinhold* contends that morality and imputability of actions can only be conceived if one presupposes freedom of the will that is distinct from the self-activity of *reason* and from the striving of *desire*" (GO, 157). Schelling argues that the reason for this dispute lies in their different conceptions of the will. In order to overcome the dispute, he interprets both conceptions of will as *aspects* of the same thing: "Here the question arises whether the object (the will) makes such a double perspective possible" (GO, 161). Schelling therefore further analyzes the concept of a *finite* will "from the standpoint of consciousness" (GO, 163): "The will, insofar it appears, must necessarily appear as choice," since "the character of the finite spirit consists in eternally appearing to itself, being its own object, becoming empirical for itself." (GO, 162) Therefore, the capacity of choice can be understood in terms of volitional self-consciousness or as a *reflection* of the will. Freedom of choice, as the appearance of the absolute will, however, is not to be confused with mere illusion—it has its own structure, as Schelling emphasizes: "This appearance of the absolute will is only true freedom" (GO, 276). Against Kant, who in his late *Metaphysics of Morals* had argued that freedom of the will cannot be "defined" (MM, AA VI, 226) as choice, Schelling argues that "the faculty of choice can be explained as *the absolute will restrained by finitude*" (GO, 167).

The restraints of finitude mean that the will as choice is confronted with alternate possibilities of action and needs to determine itself to an action. Choice is therefore an aspect of the pure or absolute will. It is the manifestation of finite individuals that are inclined both by the absolute will (or, in Kant's words, pure practical reason) and empirical motives. As such, the capacity of choice is a reflective capacity "floating between the subjectivity and objectivity of willing [...] or *what determines itself in a second potency*" (STI, 277). The capacity of choice obtains an intermediate and reflective position with regard to the subjective-empirical and the objective-pure will:

If I reflect merely on the objective activity as such, then there is only natural necessity in the I; if I reflect merely on the subjective, then there is only an objective willing, which has, according to its nature, no other object than self-determination as such. If I reflect, finally, on the subjective and objective activity of determination that is above both, then there is choice in the I, and so freedom of the will (STI, 277).

This allows Schelling to avoid both empirical determinism and intelligible fatalism, and to explain the dilemma from the perspective of different reflections of the will. Empirical determinism, as Schelling puts it, “denies freedom at all,” whereas intelligible fatalism “merely focuses on pure reason, that is an activity that is only directed towards self-determination (so that we are forced to assume that reason remains for no reasons silent in all actions that are opposed to it, whereby all freedom of the will is suspended)” (STI, 277).

6 Freedom as “the capacity for good and evil”

Schelling’s *Freedom Essay* has, in the wake of Heidegger, often been interpreted as a “metaphysics of evil,” according to which “evil is metaphysically necessary” (GA 42, 277). Furthermore, Heidegger argues that individual freedom “does not play any role” (GA 42, 15) in Schelling’s philosophy. Against Heidegger’s thesis I shall argue that it is not the necessity of evil but its *possibility* that Schelling attempts to fathom in his *Freedom Essay*. It is, as Schelling puts it, “the point of most profound difficulty in the entire doctrine of freedom.” Schelling stresses that “the real and vital concept” of freedom “is the capacity for good and evil” and that it concerns the individual and personal will (FS, 352). It is, however, difficult to isolate his systematic contribution to the autonomy problem in the thicket of different threads of the *Freedom Essay*—including the theodicy problem, the problem of creation, and the problem of system. With the exception of the title and the eminent meaning of the concept of evil, Schelling’s *Freedom Essay* seems to have little in common with the specifically post-Kantian debate on freedom. All too often, Schelling’s metaphors, such as “ground,” “light,” “darkness,” or “yearning”—give rather the impression of a ‘metaphysical’ or ‘romantic’ veiling in the bad sense of the word than a clear conceptual definition of the concept of freedom. Schelling argues, however, that the concept of freedom “must not be simply a subordinate or subsidiary concept, but one of the system’s ruling centerpoints” (FE, 336) (see Noller 2015, 310–344).

In what follows, Schelling’s conception of the individual will shall therefore serve as a common thread. Since Schelling sees freedom as a “living” capacity, his various vitalistic metaphors shall be understood as expressions and phenomenological indications of personal freedom, i.e. as determinations of such an individual will. Contrary to Kant’s thesis of the “inscrutability” of freedom for evil and Reinhold’s concept of an opaque “basic ability” (“Grundvermögen”), Schelling attempts to further shed light on the problem of evil choice: “how in each individual the decision for good or evil might now proceed—this is still shrouded in complete darkness and seems to demand a specific investigation” (FE, 352).

In his *Freedom Essay*, Schelling attempted to consistently think the possibility and effectiveness of individual choice for evil (i.e. the *capacity* of evil) and therefore left the framework of the transcendental idealism of his early writings.⁹ The task of his *Freedom Essay* is not to overcome, but to supplement idealism by transforming it by means of a kind of vital realism¹⁰:

Still, idealism itself, no matter how high it has taken us in this respect, and as certain as it is that we have it to thank for the first complete concept of formal freedom, is yet nothing less than a completed system for itself, and it leaves us no guidance in the doctrine of freedom as soon as we wish to enter into what is more exact and decisive. (FE, 351)

Freedom as autonomy of reason, as Kant developed it, is just a “formal concept.” Such a concept of freedom is closely related to the individual person or “personality,” as Schelling calls it in his *Freedom Essay* by an alteration of Kant’s terminology. In clear opposition to Kant’s conception of personality and also to his own earlier theory, which had the “destruction” (I, 128) of personality as its goal, it is now the finite, individual person whose specific freedom and imputability Schelling takes into account. Schelling’s freedom project is therefore a realistic transformation of the (post-)Kantian idealistic concept of rational self-determination. Schelling indeed acknowledges the merits in the concept of rational autonomy, but with regard to the autonomy problem seeks to supplement it systematically by including a natural-philosophical dimension: “Idealism is the soul of philosophy; realism is the body; only both together can constitute a living whole” (FE, 356). Idealism needs as its basis “a living realism” for otherwise it is an “empty and abstract system” (FE, 356).

Schelling develops the concept of the individual will as a faculty which, although related to reason and its universal dimension of the moral law, is clearly distinguished from it. In addition to the altered concept of person and will, Schelling’s realistic transformation of the Kantian idealistic doctrine of autonomy consists in a different conception of nature, which is not anymore, as it was

⁹ See Kosch 2006, 88: “[M]orally responsible agents are capable of good *and* evil, and accounts of freedom that reduce freedom to rational self-determination cannot account for this. By 1809, Schelling had come to the conclusion that the investigation into the conditions of possibility of rational self-determination could shed no further light on the question of the conditions of possibility of moral agency in a more general sense, because it could shed no light at all on the possibility of moral evil.”

¹⁰ Here I do not fully agree with Gardner (2017), who argues that “the *Freiheitsschrift* neither breaks with the original concerns of German Idealism nor leads it astray” (134). Rather, I follow Kosch (2006, 88), who argues that “[t]he systematic ambitions of German idealism [...] relied on the reduction of human freedom to rational self-determination.”

in Kant, “the epitome of all phenomena” (CpR, A 114), but rather its ground or basis. This “true philosophy of nature” is precisely the combination of idealistic and realistic moments for a concept of individual freedom that Schelling demands for his theory of freedom.

7 Schelling’s real compatibilism

By his systematic inclusion of nature, Schelling aims for a compatibilist concept of freedom. Being a natural entity does not rule out our freedom of the will: “Every organic individual exists, as something that has become, only through another, and in this respect is dependent according to its becoming but by no means according to its Being.” (FE, 346) This natural basis and dependence is not to be misunderstood as heteronomy: “dependence does not abolish independence, it does not even abolish freedom [...] dependence does not say what the dependent is or is not.” (FE, 346) Within the framework of his philosophy of nature, Schelling does not regard the naturally arising existences as a fatalistic context of necessity but interprets it as a productive grounding context that makes freedom possible and which, despite its determination, permits alternative possibilities of choice (see Noller 2015, 313–325).

Schelling’s analysis of individual existence based on nature is grounded in his theory of the will. Schelling summarizes his voluntaristic approach to the problem of freedom in what at first glance appears to be a dark sentence: “Will is primal Being [Ursein].” The fundamental concept with which the world can be understood as a place of freedom is the concept of will: “In the final and highest judgment, there is no other Being than will.” (FE, 351) The general characteristic of this will is its “self-affirmation.” This self-affirmation can be understood in such a way that the dimension of self-reflexivity belongs to willing itself. Schelling characterizes this rudimentary structure of basic self-determination with predicates such as “groundlessness, eternity, independence from time,” whereby “groundlessness,” as will be shown below, must not be confused with *indifference* (FE, 350).

Since for Schelling natural existence and will are closely connected, it is now possible to analyze the free person as emerging out of nature: “The principle, to the extent that it comes from the ground and is dark, is the self-will of creatures” (FE, 362). On a basic level, the ground of existence can be understood as “feeling,” “longing,” and “desire,” in short: as a “real” principle that provides the capacity of freedom. This mere self-will of natural existence possesses a rudimentary volitional structure and, as a “blind will” or as a “will in which there is no intellect,” does not yet have self-consciousness but nonetheless constitutes

living individuality as the basis of a personal existence. Compared to Kant's locating of inclinations in the realm of heteronomy, according to Schelling "everywhere where lust and desire" can be found, "a kind of freedom" is already present (FE, 376). This rudimentary volitional structure thus only constitutes "the ground of every special life in nature." The natural basis of one's own will is in this respect only the ground for the fact that personal individuality in the sense of a free and reflexive will (of "second-order volitions") can emerge from it at all. But in the rudimentary first-order will a reflected second-order will is already implied, or as Schelling puts it: "[O]nly from the obscurity of that which is without understanding (from feeling, yearning, the sovereign [herrlich] mother of knowledge) grow luminous thoughts" (FE, 360). The first-order will therefore already contains a second-order will: "not a conscious but a divining will [ahnender Wille] whose divining is the understanding" (FS, 359).

Such first-order desires are not self-conscious insofar as they are directly action-oriented. They are not reflected, but are bound to the object of immediate desire, without opening a scope for reflection by suspension. The second-order will, however, which is grounded in the first-order desires, is an independent and perfect will, "since the understanding is really the will in will" (FE, 359). Schelling thus grasps reason in a volitionally structured way: The second-order will as mind [Geist] is essentially self-reflective. Schelling describes this reflective structure of the will as "the word of this yearning," which means a propositionally articulated structure, "something comprehensible" (FE, 361) and as "rule, order and form" (FE, 359).

The specific distinction of human freedom becomes clear in contrast to the constitution of the *animal* will. The animal will "has not yet been raised to (does not grasp) complete unity with the light" (FE, 362), as Schelling puts it. The animal is unable to reflect upon its nature. Because it is only able to form first-order desires, it is thus a mere instinctive existence (a *wanton* so to speak). The animal's own will does not rise to the status of a free individual, but serves as a special existence only as an instance of its species, which Schelling describes by means of the concept of the "general will."

The still unconscious will of the natural individual can be understood as an instinct for self-preservation, but is basically merely an instrument and means of the universal will, that is the species. The animal cannot behave freely towards the universal will but is completely determined by it. The relationship between self-will and universal will is thus *symbiotic* in the animal: the general will enables special existence in the sense of freedom of action at all, while the self-will itself is a concrete, affirmative expression of the universal will in the mode of adaptation and regeneration. The animal is merely an object of the general, it consists, so to speak, of finite generality without any individuality.

In contrast to the animal, the relationship between self-will and universal will in humans is not rigid but flexible. Man is not only his nature, but exists in a free relation to it. The rational self-will can behave freely to the general will by its self-reflection, unlike the animal, but it can do so only by means of the general will. Man is thus able to withdraw from the general will and use it as his tool and thereby reverse the original order.

How can the complex relationship between self-will and universal will in the individual person be further determined in terms of freedom? While the animal remaining in the 'ground' is able to form only first-order desires that serve the general will, the rational and individual person is able to reflect upon the first-order desires and to form higher-order volitions, with which the person can position herself regarding the demands of the general will. It is important to emphasize that personal freedom is not limited to morality alone, but manifests itself in every practice as an expression of individuality. Second-order volitions are not to be understood in the sense of a general law—the Kantian moral law—but in the sense of the use of this general will, which is entirely 'interwoven' with the self-will.

By the reflectivity of the human mind over its constituent principles of self-will and universal will, at first only negative freedom is realized: "Since it is spirit, selfhood is therefore free from both principles." (FE, 363) Self-will and universal will only incline the person, they do not, however, necessitate her. With regard to negative freedom, Schelling agrees with Reinhold, inasmuch as freedom alone is not situated in the realm of reason alone, but in the practical use of reason. Positive freedom thus consists in the deliberate reflection of the mind on the principles of general and universal will that constitute it.

By his natural-philosophical foundation of the person Schelling detaches the moment of spontaneity as causality of freedom from the capacity of pure practical reason. *Principium diiudicationis* (the moral law) and *principium executionis* (the moral feeling of respect), which according to Kant were basically moments of pure practical reason, are now referred to Schelling's internal differentiation of the individual will. Schelling transforms the moment of spontaneity and execution, which in Kant's concept of autonomy had been thought of as the rational motive of respect, into a reflective self-will. According to Schelling's conception of the general will, the *principium diiudicationis* of the moral law is no longer thought as extra-temporal and a priori, but as historically and socially situated (see Buchheim 2012, 197–199).

8 The reality of evil

After having derived the “general possibility of evil” and thus the concept of negative freedom, Schelling is, in a second step, concerned with further determining freedom in the positive sense, on the basis of an analysis of the reality of evil. For Schelling, the reality of evil is “the subject of a completely different investigation,” since “possibility does not yet include reality, and this is actually the greatest object of the question.” In contrast to Kant, who considered the reason of immoral actions as “inscrutable,” Schelling attempts to explore “how in each individual the decision for good or evil might now proceed” (FE, 381). Schelling’s aim after Kant and Reinhold is to make the foundation and freedom of an individual personal character comprehensible.

Schelling’s conception of evil gains its profile in contrast to traditional theories of privation, but also to specifically modern ideas of evil that are due to a kind of “philanthropy” and whose ultimate consequence, according to Schelling, is even the “denial of evil.” Schelling criticizes privative conceptions of evil, for the “only reason for evil” is, according to them, “sensuality or animality.” These privative conceptions follow from an understanding of freedom that is too closely bound to reason and “according to which freedom consists in the mere rule of the intelligent principle over sensual desires and tendencies, and the good comes from pure reason” (FE, 371).

Schelling’s conceptual development of freedom to evil summarizes the post-Kantian discussion regarding the problem of intelligible fatalism. According to Schelling, “the weakness or ineffectualness of the principle of understanding can indeed be a ground for the lack of good and virtuous actions, yet it cannot be a ground of positively evil ones and those adverse to virtue.” (FE, 371) If evil actions stem from sensuality of external influences, the acting person would be necessitated and not imputable for her actions so that finally “evil would have no meaning at all.” (FE, 371 f.)

Evil, therefore, does not consist in the absence of rationality, but in its skillful exploitation, which makes it an extended form of rationality compared to the good, as Schelling puts it: “The understanding yields the principle of evil” (FE, 367), and concepts like “lack” or “deprivation” are, according to Schelling, “in complete conflict with the actual nature of evil.” (FE, 368) Therefore good and evil actions arise from the same volitional structure, they do not differ in their elements, but in their combination: “The same thing that becomes evil through the will of the creature (if it tears itself completely free in order to be for itself), is in itself good” (FE, 400). In order to overcome evil and act morally, one must not suspend or suppress one’s individuality but rather *situate* it within the universal

will and the self-will, as Schelling emphasizes: “Good without active selfhood is itself inactive good” (FE, 400).

9 Freedom as volitional necessity

Schelling distinguishes his theory from two prominent, albeit false, opinions concerning freedom of the will. On the one hand freedom must not be understood as a “completely undetermined capacity” i.e. in the sense of a *libertas indifferentiae*, according to which “one or the other is chosen by two adversarial opposites, without determining reasons” for this would lead to “nothing other than the complete contingency” (FE, 381). Schelling does, however, not reject the concept of choice [Willkür] as such, but only a certain, simplified and reduced form of it, insofar it lacks any rational structure. According to Schelling, every free action must therefore be determined, and determination is fully compatible with freedom. The “system of the equilibrium of free will,” which is the concept of indifference, must not be opposed with a system of determinism or predeterminism, which indeed gives “complete justification” of the actions but “claims the empirical necessity” so that they “no longer remain within our power during the action itself.” (FE, 382) Both systems—indifferentism and predeterminism—belong, according to Schelling, to the same standpoint insofar as they are distinguished from a third position that mediates both. This third position is Schelling’s critical voluntarism. Schelling argues for a compatibilist view, according to which “necessity and freedom must be unified” (FE, 384). He thereby refers to the concept of a “higher necessity” or an “inner necessity” (FE, 384), which shall unite necessity and contingency insofar as necessity is not thought as something external and heteronomous to man, but as a form of volitional self-affirmation:

higher necessity remains unknown which is equidistant from contingency and from compulsion or external determination, which is, rather, an inner necessity springing from the essence of the acting individual itself. (FE, 382)

Schelling’s rather obscure speech of freedom as a necessity can be explained using the example of the phenomenon of weakness of will. Assuming the case, the second-order volitions that constitute the essence of the person had no influence on the first-order desires (for example, because they themselves were inconsistent). Then those first-order desires would immediately lead to an action. The person, however, would not fully identify with their desires. This is precisely what Schelling means with “inner necessity”: a harmonious relation-

ship between first-order desires and second-order volitions, of which the respective action is an authentic expression. According to Schelling, inner necessity is therefore essentially a volitional and deliberate necessity.

We can better understand Schelling’s conception of an “inner” or “higher” necessity if we relate it with Harry Frankfurt’s account of freedom of the will. Frankfurt distinguishes volitional necessity from logical or causal necessity that would make freedom of the will impossible:

A person who is subject to volitional necessity finds that he *must* act as he does. For this reason it may seem appropriate to regard situations which involve volitional necessity as providing instances of passivity. But the person in a situation of this kind generally does not construe the fact that he is subject to volitional necessity as entailing that he is passive at all. People are generally quite far from considering that volitional necessity renders them helpless bystanders to their own behavior. Indeed they may even tend to regard it as actually enhancing both their autonomy and their strength of will. (Frankfurt 1982, 264)

As Frankfurt argues, a person who is determined by volitional necessity is in a situation which differs significantly from situations of addictions or compulsions. For volitional necessity is just an expression of the result of an inner volitional formation that has come to an definite end, which is a decisive action. Volitional necessity, in other words, can be understood in terms of strength of the will or as a kind of volitional coherence by identification, which Frankfurt calls “wholeheartedness” (Frankfurt 1987, 33). Volitional necessity is fully compatible with freedom, since it is the result of a deliberate process, a process of reflecting and balancing reasons for an action, and of integrating different volitional tendencies into an unified will.

10 Conclusion: Situating freedom

What exactly is Schelling’s progress with regard to the autonomy problem after Kant and Reinhold? Schelling’s systematic contribution to the autonomy problem can be seen in the fact that he locates the capacity of choice (*Willkür*) not only in the structure of human capacities but also within nature and the world itself. While Kant had bound the capacity of choice to pure practical reason and its moral law, Schelling’s conception of choice is flexible with regard to the self-will and the general will. Unlike Reinhold, Schelling does not grasp the capacity of choice as a “basic faculty” (*Grundvermögen*). Rather, he tries to analyze it as a volitional structure that reflects both the self-will and the general will. The reason for a free decision is therefore nothing external to both principles of will, but a reflective dimension of the will itself. Since it is not pure reason based

on a general law but the person as a voluntary unity, the free decision is situated in various respects. The person can act on the basis of her own will only in relation to her nature, and by the universal will only through the external perspective or the perspective of a group of other persons. Individual freedom can therefore be defined as situated and self-positioning, which, as a rational *process* of self-formation, also involves a historical character (see Buchheim 2012, 197). According to Schelling, freedom is not merely an inner-personal harmonic state between first-order will tendencies and second-order volitions, but is realized in various contexts. The predicates of good or evil are therefore no longer limited to individual actions and maxims, but also apply to interpersonal connections as a context of freedom (a good or evil “spirit” [Geist], as Schelling puts it) that can be donated and spread over the world.

Bibliography and Abbreviations

Citations of Kant’s works refer to the volume and page number of the Academy Edition of Immanuel Kant, *Gesammelte Schriften* (Berlin: Walter de Gruyter and predecessors, 1900–).

- CPrR: *Critique of Practical Reason*. Edited and translated by M. Gregor. Cambridge, UK: Cambridge University Press, 1997, orig. 1781.
- GMM: *Groundwork of the Metaphysics of Morals*. Edited and translated by M. Gregor. Cambridge, UK: Cambridge University Press, 1997, orig. 1785.
- MM: *The Metaphysics of Morals*. Introduction, translation, and notes by Mary Gregor. Cambridge, UK: Cambridge University Press, 1991, orig. 1797.
- RBMR: *Religion within the Boundaries of Mere Reason*. In: *Religion and Rational Theology* (pp. 39–215). Edited by A. W. Wood and G. Di Giovanni and translated by G. Di Giovanni. Cambridge, UK: Cambridge University Press, 1996, orig. 1793.

Citations of Schelling’s work refer to the Historical-Critical Edition of Schelling’s Works (HKA) by the Bavarian Academy of Sciences, or to his *Entire Works* (SW), edited by Karl Friedrich August Schelling.

- FE Philosophical Investigations into the Essence of Human Freedom (Philosophische Untersuchungen über das Wesen der menschlichen Freiheit). In: SW VII. Translated and with an Introduction and Notes by Jeff Love and Johannes Schmidt. New York: SUNY Press, 2006.
- GO General Overview of the latest Philosophical Literature (Allgemeine Übersicht der neuesten philosophischen Literatur). In: HKA I, 4, pp. 58–190, ed. Wilhelm G. Jacobs/Walter Schieche. Stuttgart-Bad Cannstatt: Frommann-Holzboog, 1988.
- I Of the I as Principle of Philosophy (Vom Ich als Prinzip der Philosophie). In: Hartmut Buchner/Jörg Jantzen (eds.): HKA I, 2. Stuttgart-Bad Cannstatt: Frommann-Holzboog, 1980, pp. 69–175.

- STI System of Transcendental Idealism (System des transzendentalen Idealismus). In: HKA I, 9, pp. 23–334, ed. Harald Korten/Paul Ziche. Stuttgart-Bad Cannstatt: Frommann-Holzboog, 2005.
- Allison, Henry (1996): *Idealism and Freedom. Essays on Kant’s Theoretical and Practical Philosophy*. Cambridge, UK: Cambridge University Press.
- Bojanowski, Jochen (2006): *Kants Theorie der Freiheit. Rekonstruktion und Rehabilitierung*. Berlin/New York: De Gruyter.
- Buchheim, Thomas (2012): “Der Begriff der ‘menschlichen Freiheit’ nach Schellings ‘Freiheitsschrift’”. In: Friedrich Hermann/Dietmar Koch/Julia Peterson (eds.): *Der Anfang und das Ende aller Philosophie ist Freiheit! Schellings Philosophie in der Sicht der neueren Forschung*. Tübingen: Attempto, 187–217.
- Frankfurt, Harry G. (1982): “The Importance of What We Care About.” In: *Synthese* 53, pp. 257–272.
- Frankfurt, Harry G. (1987): “Identification and wholeheartedness.” In: Ferdinand David Schoeman (Ed.): *Responsibility, Character, and the Emotions. New Essays in Moral Psychology*. Cambridge, UK: Cambridge University Press, pp. 27–45.
- Fugate, Courtney D. (2015): “On a Supposed Solution to the Reinhold/Sidgwick Problem in Kant’s Metaphysics of Morals.” In: *European Journal of Philosophy* 23(3), pp. 349–373.
- Gardner, Sebastian (2017): “The metaphysics of human freedom: from Kant’s transcendental idealism to Schelling’s *Freiheitsschrift*.” In: *British Journal for the History of Philosophy* 25/1, pp. 133–156.
- Heidegger, Martin (1988): “*Schelling: Vom Wesen der menschlichen Freiheit* (1809).” In: *Gesamtausgabe* (GA), Bd. 42. Frankfurt/M.: Klostermann.
- Hudson, Hud (1991): “Wille, Willkür, and the Imputability of Immoral Actions.” In: *Kant-Studien* 82(2), pp. 179–196.
- Kosch, Michelle (2006): *Freedom and Reason in Kant, Schelling, and Kierkegaard*. Cambridge, UK: Cambridge University Press.
- Martin, Wayne (2017): “Fichte’s Kreuzer review and the transformation of the free will problem.” In: *European Journal of Philosophy* 6, pp. 717–729.
- Noller, Jörg (2015): *Die Bestimmung des Willens. Zum Problem individueller Freiheit im Ausgang von Kant*. Freiburg, München: Verlag Karl Alber.
- Noller, Jörg (2019a): “‘Practical Reason is not the Will’: Kant and Reinhold’s Dilemma.” In: *European Journal of Philosophy* 27, 852–864. <https://doi.org/10.1111/ejop.12448>.
- Noller, Jörg (2019b): “Reason’s Respect: A Systematic Reconstruction of Kant’s Theory of Moral Respect”. In: *SATS. Northern European Journal of Philosophy* 20(1), 1–18. <https://doi.org/10.1515/sats-2019-0012>.
- Reinhold, Karl Leonhard (2008, orig. 1792): *Briefe über die Kantische Philosophie. Zweyter Band*. Ed. by M. Bondeli. Basel: Schwabe.
- Reinhold, Karl Leonhard (1831): “*Brief an Jens Baggesen vom 28.03.1792*.” In: *Aus Jens Baggesen’s Briefwechsel mit Karl Leonhard Reinhold und Friedrich Heinrich Jacobi*. In zwei Theilen. Erster Theil. Dezember 1790 bis Januar 1795. Leipzig: Brockhaus, pp. 166–169.
- Sidgwick, Henry (1888): “The Kantian Conception of Free Will.” In: *Mind* 13, pp. 405–412.
- Ware, Owen (2019): “Freedom immediately after Kant.” In: *European Journal of Philosophy* 27, pp. 865–881. <https://doi.org/10.1111/ejop.12428>.

Daniel Wenz

Hegel's Logical Foundation of the Will: Reconciling Psychology and Social-Ontology

Abstract: In Hegel's mature philosophy the term "will" has a socio-ontological, a subjective-psychological and a practical meaning. In this paper, I will explore the theoretical framework that enables him to combine these different strains into one general structure. To that end, I focus on the relevant categorial framework Hegel develops in his *Science of Logic* – especially on the concepts of the logical drive and modality.

1 Introduction

A common definition of a free individual is someone who lives by her own rules, by the rules she chooses for herself. However, to be able to do this presupposes being able to act upon a rule. In this case, it is not sufficient that the behavior of such an individual can be interpreted or predicted by a rule. She herself must understand the rule and know what counts as acting according to it and what not. However, as a well-known argument goes,¹ a rule is always a generalization, and its application in a concrete situation is in need for another rule that determines how the rule has to be modified to fit the given context – and the same goes for this rule etc. So either one cannot act at all or has to rely on some kind of implicit knowledge or know how. In the latter case, some 'background' has to be presupposed that is encoded in the mere behavior of the individual as well as in the design of the artifacts (doors and tables) that surround her and the social institutions (friendships, banks and states) that constitute the social reality she lives in.

However, this is only one side of the problem. To say that being free means to live by one's own rules begs the question of how the free individual chooses the rules she lives by. It is easy to say that she chooses according to her own will – but what does this will consist in? To reply that her will is determined by her personal preferences is not of much help: Those preferences will consist of her more or less justified beliefs and convictions. All the same, these are in no meaningful

¹ The following paraphrases Searle 1983, 141–159.

sense chosen by her. Her preferences are a product of the natural and social world she inhabits. They are products of her experience with the world or learned and conveyed by tradition. To point to the natural side of things will not help. The desires, drives, and feelings that constitute her individuality are part of her natural setup – and this is not for her to choose. All this makes it hard to see how a free individual in the sense sketched above could exist. In what follows, I will argue that Hegel articulates a concept of the free will that takes into account the social and natural determinateness of human beings.

Hegel primarily develops the different aspects of the notion “will” in the third part of his Berlin *Encyclopedia of the Philosophical Sciences from 1830* (*Enzyklopädie der philosophischen Wissenschaften*) and in the *Elements of the Philosophy of Right* (*Grundlinien der Philosophie des Rechts*).² In the latter, the concept of the will is introduced as the foundation for the explication of the different institutions that realize the concept of freedom. This can be read as an exploration of the relation between the will as (a) an anthropological and psychological presupposition, (b) the will of an individual person and (c) the will as an institutionalized, socio-ontological entity. Roughly, (a) and (b) together constitute what Hegel calls the subjective will while everything that falls under (c) he denotes as the substantial will. The subjective will corresponds to what in the exposition above was called the free individual. The substantial will approximately equals what was called the background.

The subjective will is one of the main topics of what Hegel calls the philosophy of the subjective spirit, while the figure of the substantial will is articulated in his philosophy of the objective spirit.³ In respect to modern philosophical disciplines the former comprises a philosophical physiology and anthropology, a philosophy of embodiment and a philosophy of action and thought, while the latter articulates a social ontology.

I will argue that Hegel’s methodological approach can point us to a solution to the conundrum sketched above and that it can be better understood by reconstructing its underlying categorial framework as developed in his *Science of Logic*. The leitmotiv for this is the concept of the logical drive (*logischer Trieb*) and his take on modality.

² In accordance with the international standard I will cite the *Encyclopedia* and the *Elements* by paragraph (Ency. §; Elements §). Hegel’s *Science of Logic* and his *Phenomenology of Spirit* are cited according to the respective volumes of the *Gesammelte Werke*. For the English translations, I refer the reader to Hegel 2010 and Hegel 2018, which both provide the numbering of the *Gesammelte Werke*.

³ For the relationship of subjective and subjective spirit cf. Ency. § 385.

After a thematic introduction, I will reconstruct how the concept of the logical drive is introduced via the attempt to articulate the concept of objectivity. Against this backdrop, I will present what Hegel calls the subjective will as a methodological starting point and introduce from this perspective the concepts of a genetic and a constructed presupposition. Finally, I will dive into Hegel's theory of modality as a preparation for the reconstruction of three different interconnected models of the will that can be found in the *Encyclopedia*.

2 The concept of the subjective will as a common sense picture of the human condition

In Ency. § 399 Hegel coins the sensation (*Empfindung*) as the unconscious counterpart of the will. A few paragraphs later (Ency. § 403) the feeling (*Gefühl*) is introduced as the self-relation that is contained in the sensation (*Empfindung*). This self-relation that connects all sensations as feelings is supposed to constitute the differentiation between a unity of feelings and a somatic (*körperliche*) unity.

When Hegel at this stage writes about sensations and the unity of the body, this evokes a certain picture in us: We know sensations as discrete feelings, as a kind of non-verbal private content that we have. My toothache is my private pain, that no one else can have or experience. And it somehow is located in or on my body (my teeth) and evoked by something outside of my body. So by having a feeling, I implicitly differentiate between the world, the things outside of me that caused the feeling, the unity of my body that this feeling is mapped onto (and in this way becomes a somatic unity) and which interacts with the outside world, the feeling itself and me. This me could thus be described as the center or the unity of my feelings.

If I reflect upon this me, I reflect upon myself – but in the instance that I realize this fact I am more than just the unity of my feelings, I am an I that consciously places itself into a world and that differentiates this world from its body and this body from itself: I conceptualize myself as a thinking self, while the unity of my feelings somehow resides as a 'dark' undertone inside myself.

To think about this condition already presupposes being in a reflective mode. Because it is impossible to be a purely bodily and feeling creature and at the same time think about this fact. Endorsing this position does not commit one to deny that we have a direct connection to our feelings or an immediate sensual representation of our body. But it contradicts the idea that we have an

epistemic privileged first hand approach to these domains, even if its material is given to us directly by introspection.⁴

We can reflect on the preconditions that enable this reflective subject to reflect its internal preconditions. This takes us from an introspective investigation concerned with internal conditions to a reflective investigation of our external conditions. Every iteration of this reflective act, every reflective stage generates a new reflective subject: We find ourselves as an active being in our world, i.e. a being that is able to rationally comprehend and change its world. We find ourselves in a world that is shaped by other such beings, which allows us to understand ourselves as such a being etc. At some point this iteration of reflective stages results in a conception of the I or me as a certain kind of will. This kind of will Hegel calls the subjective will (*der subjektive Wille*) or the European concept of freedom (*Freiheit im europäischen Sinne*) (cf. Ency. § 503).⁵

This conception of the will is, as will be shown, a fairly good description of our everyday experience with ourselves. The main characteristics of this will are the internalized capacity of an independent differentiation between good and evil – a feeling for what is right and what is wrong – and that it understands itself as an end in itself. Being an end in itself implies that the determination of the will subsists in itself. Therefore, the behavior of such a subject is not an agency but an action (*Handlung*). It accepts only those acts as its own actions that are an expression of or that are caused by its own intentions.

It will be shown that the European concept of the will (i.e the concept of the subjective will) is the methodological starting point for the distinction of internal- and external preconditions. This follows from a figure Hegel develops in the categorial framework that is his Science of Logic. However, before any further explanation of this application of Hegel's logic can be given, the question has to be asked how its categories are connected to the real world.

3 Categories and the real world

So we are in the dark about the nature of philosophical problems and methods if we are in the dark about types or categories (Ryle 1938, 189).

⁴ This of course does not mean that qua human beings we have never been in such a state. Child development refutes such an idea.

⁵ The addendum "European" demarcates the historicity of the concept. This means that it presupposes a specific historio-sociological development.

According to Hegel philosophy draws its content from the conflicts that arise out of our theoretical and practical meddling with the world. These conflicts demarcate philosophical categories.⁶ In Hegel's non-logical writings and in his annotations to the *Science of Logic*⁷ categories are explored mainly in a material way. These investigations usually depart from the breaking points that emerge in the special sciences or our established everyday practices. They uncover the boundaries of the respective theory or of our semi-institutionalized habits of practical orientation.⁸

Those breaking points manifest themselves in the form of contradictions, incommensurabilities and dilemmas. To give a couple of examples: The disciplines of pure mathematics are (according to Hegel) brought to their limits when – for inner-theoretical reasons – they are forced to take into account the qualitative basis of the formal operations that define them. In the annotations of the second edition of the first part of the *Science of Logic* Hegel explores some of these cases in great depth. An infamous example are Hegel's remarks concerning the infinitesimal/differential calculus (cf. GW 21, 250–299).⁹ Along these lines Hegel's whole exposition of his *Philosophy of Nature* can be read as an investigation of the categorial transgressions in the natural sciences of his day. Breaking points that appear in the practical-social realm can be found in abundance in part VI of Hegel's *Phenomenology of Spirit*. The most prominent case is that of his interpretation of Sophocles' Antigone. In the ancient play one agent whose character represents two different codes of moral conduct gets caught in a situation where both contradict each other.¹⁰ In this instance the category error

6 It is very interesting to compare this approach with Ryle's Tanner Lectures (Ryle 1953).

7 Angelica Nuzzo interprets Hegel's whole dialectical logic as an attempt to deal with such breaking points (cf. Nuzzo 2006).

8 Especially for his *Phenomenology of Spirit* cf. Heinrichs 1974, 71–89.

9 A lot of literature is dedicated to Hegel's investigation of the calculus: Wolff 1986, Moretto 1993, Paterson 1997, Stekeler-Weithofer 2005. The calculus was (and in many ways still is) the mathematical tool of choice to approach natural phenomena. It originates in a type of mathematical thinking that was more experimental and focused on applicability than on rigor (cf. Grabiner 1981). This makes it a good subject to study the mixing of different types of categories and categorial transgressions. Hegel wrote his annotations in a transitional stage. In his time, the basic concepts of the calculus only began to receive their exact definitions. This development was brought forward mainly through the cutting-edge research of Cauchy that Hegel also takes into account (cf. Wolff 1986). The calculus that we know today is due to Weierstrass' systematization of Cauchy's work (cf. Volkert 1987).

10 Antigone, the main protagonist, has – according to the divine law (the law of the family) – the duty to bury her dead brother, who is an enemy of the state she lives in and to whom by state law a burial has to be denied. This small part of Hegel's *Phenomenology* has spawned research literature in abundance. Just to give a few examples: Mills 1986, Kain 2002, Chanter 2010.

(which is supposed to happen on a purely practical level) is felt by the protagonist as a pain (pathos) that gives her insight into the social structure of the world she lives in (cf. GW 9, 255–260).

One of the tasks of philosophy is to carve out these categories and use them as a means of philosophical analysis. The targets of this analysis are the same instances that those categories are extrapolated from (or that can be identified as falling under them). The basic idea is to criticize things by their own standard: Not to judge them according to a theory that is external to them but by exposing their inner deficiencies and so further their internal development.

Hegel's *Science of Logic* is an attempt to systematically develop and ground such categories according to their own inner logic. This implies that for him the historical genesis of the categories does not prove (or disprove) their validity. Their validity feeds from conforming to criteria that are set by their own systematic presentation. Although their excavation takes place as a historical process (as the progress of philosophy in time, understood as an ongoing intellectual discourse), it is supposed to be a purely conceptual development. In this sense, the validity and meaning of the categories are taken to be bound exclusively to the internal positings of the philosophical discourse.

Roughly, the leading principle of this (historical) development is (according to Hegel) a kind of conceptual necessity. The introduction of a specific category conceptually presupposes the ability to thematize certain other categories. Here are some examples: It is not possible to handle the category of substance without being able to handle the category of accident (cf. GW 11, 94–396). The category of interdependency presupposes the category of dependency (cf. GW 11, 407–409). There has to be a theoretical approach to the category of being before the difference between essence and mere appearance can be introduced (cf. GW 11, 241–243).

Although the validity of the categories cannot be reduced to their genesis, there are interdependencies between the historical-contingent and the inner theoretical developments of the categories. Hegel himself stresses this constantly: Making certain key-categories and their interdependencies explicit presupposes some concrete developments taking place in the real world (cf. GW 14.1, 16–18). Particularly those technical inventions and discoveries in the special sciences and those historical developments that establish new social realities which cause a crisis are prone to uncover new conceptual resources (cf. GW 21, 10–12). By doing so, they lead the way to new practical and theoretical perspectives that may abolish established paradigms or uncover formerly hidden frictions. In a nutshell, Hegel claims that the theoretical exploration of the philosophical categories is dependent on certain configurations of the human

lifeworld, without claiming that their validity is restricted to specific lifeworldly configurations.

Against this backdrop, Hegel establishes three ways to talk about the structure of the categories: (i) A *Science of Logic* thematizes and systematizes the categories according to their internal criteria and interdependencies. (ii) A *History of Philosophy* organizes the philosophies that have arisen in the history of the world according to the categories that they express. (iii) A *Philosophy of History* as a reconstruction of the crises that appear in the social and cultural realm according to the logical development of the categories.

These different options mirror the relationship between the genesis and the validity of the categories. The execution of (i) depends genetically on what is developed in (ii) and (iii): The articulation of a *Science of Logic* presupposes that certain key-concepts already have been developed in the course of the history of philosophical reflection (ii). To conduct (ii) presupposes that certain socio-cultural crises (iii) that can be the target of a philosophical reflection have already occurred. Then again (ii) and (iii) are dependent on what is thematized in (i): *Post hoc* it can be stated that the different philosophical theories of (ii) had to appear in a certain order because a specific category can only be thematized after its logical predecessors have been introduced. Finally, the crises that appear in the social-cultural realm can only appear because they are composed in a way that renders it possible for individual and collective actions to produce practical contradictions. These contradictions can only emerge because they violate the scope of application of different categories. In this sense, a crisis is a category error. The concept of the logical drive explicates the structure of this kind of productive contradiction. It is the model for the concept of the subjective will and its special methodological status.

4 Pro- and Retroactive self-preservation: The logical drive of the subjective will

In his *Science of Logic* Hegel introduces what he calls the logical drive (*logischer Trieb*). The logical drive is a form of productive self-contradiction. It arises when the construction of a theoretical framework generates facts that contradict its foundational postulates. In such a case, the incoherent framework has to be stabilized by the articulation of a wider context. The content of the basic concepts for the construction of this new context is determined by the task to introduce a new categorial distinction that re-contextualizes the conflicting terms, thereby introducing not only a new conceptual domain but also a contradistinction

that determines the conditions under which each of the two conflicting concepts are in force.¹¹ For the main part of the *Science of Logic* the concept of the logical drive as a method of concept formation is applied only implicitly. Hegel derives its explicit form at the end of his investigation of the concept of objectivity.

Under the title “Objectivity” Hegel develops a conceptual framework according to which a certain kind of object is determined solely through the relations to other objects of the same kind (cf. GW 12, 127–172). The two main principles of this conceptual framework are the principle of insularity and the principle of relational symmetry. The concept of insularity implies that every object in this domain is determined only by other objects of the same kind. The principle of relational symmetry implies that the relations between the objects convey the same properties to each object. A conflict arises if this setup is confronted with a kind of object that depends on asymmetric relations. Asymmetric relations are those that ascribe different properties to their relata and in this way determine them as different kinds of objects. This is the case within the object-domain of teleology regarding the relation of being a means to an end.¹²

The objects in the domain of teleology are determined through an inter-relational network that determines them either as means or as ends. This setup therefore violates the principle of relational symmetry. An obvious way out of this would be to abolish the strictness of the disjunction by establishing that every object is a means as well as an end. A prerequisite for this would be the introduction of an external perspective that articulates a contextual criterion that establishes when the same object has to be taken as a means or as an end. But this option is ruled out by the principle of insularity: It prohibits any kind of invocation of an external perspective.

Hegel’s solution (not only to this specific case but to problem of asymmetric relations in the domains of objectivity in general) consists in the construction of the concept of an object as a self-generating inside-outside distinction: A self-related process that generates itself by procedurally constructing and dissolving the internal-external distinction with itself as its center. Hegel calls this structure

11 This is the general method of the constructive introduction of new concepts within the *Science of Logic*. Although the authors do not refer to the concept of the logical drive, Kreis 2015 and Rosen 2013 give a very lucid exposition of this structure. The pioneering Henrich 1967a and Iber 1990 famously highlight the importance of Hegel’s treatment of the concepts of reflection (*reflexion*) in GW 11, 249–257 for this topic.

12 This is already the case with the two previous domains Hegel introduces under the titles “mechanism” and “chemistry.” The domain of chemistry was introduced as a solution to the conflicts that arose under the mechanistic domain and the domain of teleology as a solution to the conflicts that arose under the domain of chemism (cf. GW 12, 224, 233–235).

the *Idea* and its first incarnation the *Idea of Life* (cf. GW 12, 176–177). The concept of the idea is the explicit articulation of the concept of the logical drive. Up to this point in the *Science of Logic* the logical drive was an operative concept, the method or force that implicitly governed the development of the categories.¹³ Now it is explicitly brought up in the *Science of Logic*.¹⁴ The remaining part of Hegel's logic is dedicated to articulate the stability-conditions of this structure – i.e. the conditions of self-preservation of the logical drive.

In retrospect the articulation of these stability-conditions comprises the whole development of the *Science of Logic*. Categorical structures that precede the explicit introduction of the logical drive have to be reconstructed as its *genetic presuppositions*. The structures that stabilize and proceed its introduction have to be constructed with it as its ground. They are, so to speak, the presuppositions the logical drive constructs itself – its *constructed presuppositions*.¹⁵ I propose that in the *Philosophy of Spirit* of the *Encyclopedia*, the subjective will (i.e. the European concept of freedom) is an instantiation of this category and can therefore be read as such a point of reference for the genetic/constructive distinction. The general idea is this: The primary genetic presuppositions are the constituents of the psychological and somatic setup of the subjective will. The constructed presuppositions are those external circumstances that constitute the subjective will as a social entity. The latter are actively constructed insofar as they are generated by the engagement of the subjective will with the social world. The primary genetic and the constructed presuppositions are interdependent. The primary genetic presuppositions can only be the target of a theoretical investigation if we understand them as the elements that constitute the unity of the subjective will as an active social entity.¹⁶ To phrase this more emphatic: They only exist as genetic presuppositions of this unity. Before the subjects engages as a social entity its genetic presuppositions will be something else.¹⁷ Besides the primary, there are other kinds of genetic presuppositions. The specific social

13 Most insightful in this respect is Hegel's treatment of the concept of *contradiction*, especially in the third annotation (cf. GW 11, 286–290).

14 For the operative/thematic distinction concerning the employment of the categories cf. Henrich 1978, Richli 1982.

15 Hegel himself sees the transition from the objective (the *Doctrine of Being* and the *Doctrine of Essence*) to the subjective (the *Doctrine of the Notion*) part of the *Science of Logic* as such a point where he calls the objective part of his Logic the genesis of the notion (cf. GW 12, 11–14).

16 This mirrors the fact that from the point of view of a concrete subject, they are only accessible as its own private sensations, feelings, etc.

17 Cf. for the general outline of this idea the Hegel inspired approach in McDowell 1994, 109–126.

context that a subject is brought up in but does not understand is an instance of a social genetic presupposition.

Talking about the constitution of such a being therefore takes place within the theoretic boundaries of social ontology. Roughly, what we now call social ontology Hegel calls objective spirit. Everything that is discussed before this stage takes place in the theoretical domain of subjective spirit. This domain comprises the philosophical disciplines of anthropology, a theory of consciousness, psychology and a physiologically grounded somatic phenomenology. The objects of these disciplines are only accessible via a reconstruction with the subjective will as its point of reference. Everything that succeeds the subjective will is accessible via a construction. The reconstruction interprets its phenomena in the context of the subjective will, while the construction interprets the concept of the subjective will in a wider context. Both the reconstruction and the construction can be seen as explicating its presuppositions – but in very different ways. While the reconstructed presuppositions can be called genetic presuppositions, the constructed presuppositions are meant to stabilize the subjective will.

5 Objective modality as a constructive principle of reality

The preliminary sketch of the subjective will given above describes a unity that plays two roles at once. In respect to the elements of the genetic presupposition, this unity provides a context within which they are individuated as such elements. At the same time, this unity is supposed to be composed of those elements. Although this description obviously involves a circle, the structure it describes is in general not problematic. A guitar is made out of different parts of wood and steel or nylon. And while it is true that a certain part of it only exists as a bridge in respect to the function it fulfills as part of the guitar, it still has its independent existence as a piece of wood. This is unproblematic because two different frameworks of description are in play: One provides a material (wood and steel), the other a functional (bridge, string) description. We, from an external point of view, can combine both frameworks by making identity claims between them (i.e. by postulating cross-sortal identities).¹⁸

This way to defuse the circular relationship between the unity of the subjective will and its genetic presuppositions presupposes two things: That there is an external point of view from which those (cross-sortal) identity claims can be

¹⁸ For a Hegel inspired take on cross-sortal identity claims cf. Brandom 2015, 216–235.

made and that we are entitled to make these claims. The problem is: The thing that is supposed to be generated in the case of the subjective will is the external point of view that allows to set these identities. It is the perspective from which different discrete things (sensations, body parts, etc.) are mapped onto each other.

In respect to Hegel's notion of objectivity, which provides the stage for the development of the concept of the logical drive, it is the principle of insularity which prohibits the invocation of an external point of view. According to Hegel this is why the notion of objectivity is internally linked to the notion that things are in respect to their specific determinations independent of any external description – that the world is as it is independent of how we describe it.

In physics this feature of objectivity is captured by the concept of an independent and in this sense absolute inertial reference frame, an idea that is carried over from its Galilean origins through Newtonian dynamics (from where Hegel picked it up)¹⁹ to the modern theory of relativity.²⁰ The main idea is not to exorcize the concept of different perspectives altogether but to reconstruct the point of view from which those different perspectives are articulated and are related to each other within the same domain and as the articulation of this domain. This is another formulation of the general idea of an internal self-differentiation.

To apply this strategy to the philosophy of mind and action is not so outlandish at it seems at first sight: Every theory that is designed to naturalize epistemology is an instance of it.²¹ It is also the natural stance that we take on our own position in the world – we conceive ourselves as beings that are part of the world we cognize and manipulate.

From a Hegelian standpoint, the general question that derives from these different approaches in the philosophy of mind and action, epistemology and physics is this: How can a framework be composed out of elements that depend in what they are exclusively on the determinations that are set through this same framework? Hegel tries to solve this problem before he develops his concept of objectivity and the logical drive via a logical analysis of the terms “mediated”

¹⁹ Cf. Ihming 1989, 29–152, 173–175.

²⁰ Cf. Wandschneider 1982.

²¹ Positions that take their point of departure from Quines' conception of a naturalized epistemology (Quine 1969, Chuchland 1987, Kornblith 2002) or Sellars' completed Scientific Image (especially as endorsed in his later writings, cf. Sellars 1981) would be candidates or at least in need for such a purely descriptive domain. The structure Hegel is interested in is more general and although certain varieties of naturalism can be used to illustrate it, its application does not commit one to a naturalistic position.

and “immediate.” This analysis reaches its peak in his treatment of the concept of “actuality” (*Wirklichkeit*). In a nutshell, Hegel calls an actuality a domain that carries the reason or cause for its own structure in itself (cf. GW 11, 375, 380–381). This notion of actuality uncovers the basic presuppositions of his concept of the logical drive.

Heavily influenced by Spinoza and Leibniz, Hegel realizes the exposition of this concept of actuality as the development of the concepts of modality.²² Mainly, he develops the notions of formal, real and absolute possibility, actuality and necessity. It is important to notice that Hegel here uses the term “actuality” for the whole structure he develops in GW 11, 380–393 as well as for a sub-part of this structure, the modal category of actuality. In the following, I will restrict myself to highlighting only those parts of Hegel’s complex exposition that are relevant to the topic of this paper.²³

First, let me introduce some preliminaries:²⁴ A possibility denotes something that is defined through conditions that have to be met to make it appear. Looking at it from another angle, this means that a possibility defines a certain context, a certain state of affairs: the one that is the case when the possibility (the thing²⁵ that is possible) appears. In this sense the concept of possibility is realized in the relevant context as well as in the instance this context brings about. Relative to the concept of possibility, actuality can be introduced as one or two or all of the

22 Hegel refers to Leibniz’ concepts of appetite and monad and the substance/attribute/modus structure and the concept of the conatus Spinoza articulates in his *Ethics* (GW 11, 376–379). To just roughly summarize Hegel’s reading of Spinoza: Ontologically there is only one thing, the substance. The substance is something that is comprehensible (and differentiated) only under a (one of many possible) predicative scheme(s) – an attribute. The differentiations that can be made (or that take place) under an attribute are called modi. As there is only one substance, the differentiations initiated by the different attributes and the application of the attributes in general cannot come from the outside. They have to be conducted by a specific modus under a specific attribute. According to Hegel, this is the attribute of thought and the relevant modus is the thinking subject. Actuality as a whole is this process of internal self-differentiation and the concepts of modality (actuality, necessity, possibility, and contingency) articulate its domain (cf. for this reading also Karl Ludwig Michelet’s reconstruction of Hegel’s lectures on the History of Philosophy in Hegel 1986, 157–197).

23 There is a lot of literature on Hegel’s concept of modality that provides quite different readings. For some examples of this diversity cf. Henrich 1967b, Hoffmann 1991, Stekeler-Weithofer 1992, Houlgate 1995, Lampert 2005. A recent approach that is heavily influenced by Hegel’s exposition in the *Phenomenology of Spirit* can be found in the writings of Robert Brandom. Cf. Brandom 2008, 117–175, Brandom 2015, 174–215.

24 For a more detailed approach cf. Wenz 2016.

25 This thing can itself be a state of affairs.

things that make up the state of affairs that make a possibility occur – considered in abstraction from this possibility.

The possibility that appears is also an actuality – in the sense that it is part of another state of affairs that makes another possibility appear etc. Also, every actuality is a possibility – in the sense there is a certain state of affairs that makes it appear.²⁶

This interrelationship between the concepts of possibility can be read in the following way: A possibility is the structure of a state of affairs, but thought of as a unity, realized in a single instance. To use a more picturesque phrase: A possibility is a node in a network of state of affairs whereby each state is itself a possibility in its own right. On the one hand an actuality is an appearing possibility. On the other hand it is the multidimensional network of state of affairs, ordered according to the possibilities. This picture enables us to give a preliminary definition of necessity: Necessity denotes the relations that consist between the different unities such a network constitutes.

Based on the dual role possibility and actuality play in respect to this multidimensional network, two different statuses can be attributed to them: If a possibility is taken as a unity, then it has the status of an immediacy (*ein Unmittelbares*). If it is taken in respect to the state of affairs it is associated with, it has the status of a mediated (*ein Vermitteltes*)²⁷. The same goes for actuality: As an element of the state of affairs an actuality has the status of an immediate – but taken as an appearing possibility it has the status of a mediated (*ein Vermitteltes*).

Hegel modifies the modalities with the adjectives “formal,” “real” and “absolute.”²⁸ “Real” can be roughly translated with “material” or “content dependent.” In this sense, a real possibility is a principle of individuation that articulates a discrete material content. A real actuality is either this discrete material content or that what has been divided to individuate this content. An approximation to “formal” is the complementary term, “not content dependent.”

²⁶ All this talk about “something that also can be seen as something else” is of course just an abbreviation. The rationale of Hegel's construction is much more intricate. The first step is to view modality as articulating the structure of actuality which means to take actuality as something mediated (*vermittelt*). The second step is to reflect upon the presuppositions of this structure which is actuality taken as something that is immediately (*unmittelbar*) given. The structure Hegel derives from his analysis of modality can be read as the further analysis of the concepts of positing and presupposing Hegel gives in GW 11, 249–257. Cf. Henrich 1967b.

²⁷ I adopt Hegel's idiosyncratic use of language and use “immediacy” “mediated” as noun phrases as well as adjectives. To use them as noun phrases is meant to highlight that the functional roles those terms denote themselves are the topic of the discussion.

²⁸ For Hegel's analysis of the concepts material and formal cf. GW 11, 294–302.

Again, a formal possibility is a principle of individuation, but now of a content-free unity. A formal actuality is either this unity or a structure that has been ordered to individuate this unity.

As the characterization of “formal” as not-content dependent suggests, “real” and “formal” are interdependent concepts. To explain this, let’s start with a real actuality as our point of departure. Such an actuality is thought of as somehow content dependent but not in itself articulated. It is so to speak not a discrete but a continuous reality. In its final stage it should be divided into different discrete actualities via the principle of individuation that we coined “real possibility.” But for now, this does not seem possible, because it would presuppose that there are already discrete actualities that we can construe collectively as a real possibility. Now imagine that we take a formal actuality, a not content dependent order and somehow (how exactly we don’t know yet) convey its structure onto the continuous content of the real actuality.

From the perspective of the real actuality, the resulting divisions would seem arbitrary: They are just formal unities and contingent in respect to their material content.²⁹ But according to the structure of the formal actuality, each such unity is articulated and therefore necessitated by a formal possibility – they are internally validated in respect to this structure.

Now, the formal possibility distinguishes within the real actuality a collection of different parts that bring forth a discrete unity. This is a formal unity, but as a part of the real actual it also has a content. This content is inarticulate, i.e. it has no internal structure. According to the terminology coined earlier, we can call such a content or real actuality an immediate actuality. If we now take this unity as our point of departure we can reverse this procedure and differentiate the real actuality into those real actuals that collectively constitute this content as a unity. These real actuals are supposed to have some kind of material content because they reside in the domain of a real actuality. But for now we don’t know how this determination is articulated, which means that they also have to be taken as immediate actualities. The punchline is that together they articulate a content, and that the unity of this content is not the real actual that the formal unity denotes, but the formal unity itself. So a mediated real actual is the unity of a formal unity and the collection of immediate real actuals that constitute the real possibility that the formal unity denotes.

²⁹ To reduce the complexity of my exposition, I omit any further discussion of Hegel’s concept of contingency. As it deals with the prerequisites of the incorporation of new experiences, it plays an important part in Hegel’s take on epistemology and theory of science (cf. Henrich 1967a).

Every one of the immediate real actuals of the collection that defines the articulate content has this kind of structure, i.e. it is the unity of a formal unity and the collection of immediate real actuals that constitute the real possibility that the formal unity denotes.

The formal structure we took as our point of departure is generated via the mediated relation between different domains of real actualities. So, the formal unity consists in the mediation of two or more different actualities. At the same time, those real actualities are generated through this mediation. Finally, the term “absolute necessity” just sums up this whole structure as a self-generating and self-determining process.

With the concepts of modality and the logical drive behind us, we can now try to distinguish the three main instances of the concept of the will in the *Encyclopedia*: Again, we will start with sensation.

6 The genetic presuppositions of the subjective will

As mentioned before in part 1 above, Hegel coins the sensation (*Empfindung*) as the unconscious counterpart of the will (Ency. § 399). Systematically, every sensation expresses a unity that brings different sensations together. Hegel calls this unity the “natural soul” and describes the sensations as passive judgments³⁰ that are bound together in a non-systematical and purely formal way.

That the unity of the different sensations is a formal unity does not mean that the particular sensations have no content but that the structure of this unity is not determined through these contents. Its unity is so to speak externally imposed, it is only caused through the unity of the body as a natural organism: The natural organism is the genetic presupposition of the unity of the different sensations. This imposed unity makes a systematic articulation of the relationships between the particular sensations possible – a process that results in a real/material unity, i.e. a unity that is articulated by the material content of the particular sensations. This unity is then transferred back to the body, thereby assigning to the contents a somatic topos, i.e. differentiating where or with what

³⁰ In this context the term “judgment” must not be understood in its everyday sense but as a unity that constitutes a differentiation (cf. GW 12, 53). It is one of the early instantiations of a self-constructing unity that follows from Hegel's theory of modality.

the sensation is felt.³¹ This generates a somatic unity, transforming the unity of the body from a genetic to a constructed presupposition.³² This again has consequences for the unity of the different sensations: It is now a real unity, a unity that is determined through the contentful relations that the somatic unity generates. According to Hegel this has to be conceived as an ongoing process, a constant loop. It is the unity of this whole process that Hegel in the anthropology section of the *Encyclopedia* calls the “natural soul” (cf. *Ency.* §§ 391–402).

According to this conceptualization, the unity of the sensations and the unity of the body are two interdependent aspects of the same structure: The unity of the sensations is an abstraction from the unity of the body and in order to identify the body as a whole the unity in sensation has to be presupposed. This is an instance of two different domains, two real actualities that are mediated through a formal unity and generated through this mediation. The differentiation of these two domains is constitutive for the structure of consciousness and self-consciousness that is explicated in the phenomenology section of Hegel’s exposition of the subjective spirit. This constitutive differentiation links the unity of sensation to its conscious counterpart, the will.

The will in its first incarnation is the topic of Hegel’s psychology. It is, as the active counterpart to sensation, the practical sensation (cf. *Ency.* § 471). This results in another loop that connects the natural soul and the will as the practical sensation: The will as practical sensation presupposes the reconstruction and thereby systematization of the unsystematized judgments of sensation. This is the function of the different iterations of what Hegel calls the theoretic spirit (cf. *Ency.* §§ 445–468). Although Hegel’s exposition of the theoretic spirit is very brief, it also is immensely complex. It comprises the interaction between representation, productive and re-productive imagination, memory, language and thought, all centered around the notion of understanding (*Erkennen*). This is not the place to give an interpretation of Hegel’s intricate approach to unify these concepts. What matters here is that for him understanding is not the integration of a given content in an already existing system of judgments (in a web of beliefs) or the reproduction of a given order the mind as a mirror of nature. “Understanding” is the construction of a system of judgments whereby the construction consists of the justification of its content. In the domain of understanding nothing is taken as a given (as in sensation), instead it has the (modally articu-

31 Instead of a sensation of pain that can be located somewhere on my body I just have a sore hand or an aching foot.

32 The difference between the unity of the body and the somatic unity resembles the phenomenological distinction between body (*Körper*) and phenomenological body (*Leib /corps propre*) as introduced by Husserl and Merleau-Ponty (cf. Husserl 1982, Merleau-Ponty 2002).

lated) status of something immediate or mediated. Neither the content nor its order and structure are given and no criteria of rationality that guides the process of justification is presupposed, rather they are all constructed.

The unity of the system of judgments whose construction is the activity of the understanding is mirrored by the practical sensation (*praktisches Gefühl*). The practical sensation is part of what Hegel calls the practical spirit (cf. Ency. §§ 469–480) – the counterpart of the theoretical spirit. The practical sensation expresses this unity as the sensation of a “situation” (cf. Ency. §§ 471–472). This situation is not the picture of the place of someone in his world but more like the somatic unity of the body, which is as such not identical with the physiological setup of a natural being. The situation of practical sensation is the judgment (the self-determination via an internal differentiation) of this unity as either pleasant or unpleasant.

It is important to notice that this kind of is-ought difference is not the difference between the way the world is vs. the way someone thinks the world ought to be. It is an internal judgment that is interdependent with the status of coherence and incoherence of theoretical spirit. A “drive” (*Trieb*)³³ highlights one of the parts of this situation by relating it to the unity of this situation (cf. Ency. §§ 473–475). This is an instantiation of the logical drive in its implicit form. The connection it establishes between part and unity always consists in a re-ordering or re-construction of this unity.

According to this picture, the unity of the situation is generated through the ongoing process of the determination of drives. In this context “Interest” (*Interesse*) is defined as the object of a drive (cf. Ency. § 475).³⁴ It is important that at this stage we are talking about an internal process: The object of the drive is not something in the world that is desired by someone but a self-objectification. The unity of practical sensation is the object that is expressed in this manner by focusing itself upon a selection of specific drives. This selection Hegel calls “Willkür” (*Willkür* = *Wille* + *Kür* = election of the will) (cf. Ency. §§ 476–478).

Prima facie, “Willkür” means “arbitrariness” as there is no criteria regarding this selection. But *Willkür* also is the election of the will as the unity of this selection. In a certain sense, this means that the will chooses itself. But on this level this does not mean that the will itself provides reasons for this selection. As the *Willkür* is part of an ongoing process, the will as the specific setup of the practical sensation is the relative and changing ground for every selection.

³³ The German word “*Trieb*” can be translated as “drive” or “passion.”

³⁴ This is the reason why the drive in this instantiation emerges as many drives: Its objectification is realized through different objects.

The unity of the will is, as the unity of sensation (taken externally as the unity of the body) before, only a formal unity.

When this process is objectified and taken as a whole as the expression of the will Hegel calls it “happiness” (*Glückseligkeit*) (cf. Ency. §§ 479–480). On another level, happiness – in a Kantian vein – has the function of a regulative principle (cf. Kant AA 4, 124–132). But in the context of Hegel’s anthropology and psychology, the will is just the unity of practical sensation as structured by passion and organized by interest. Thereby the special kind of unity of the will as practical spirit is interdependent with the unity of theoretical spirit.

7 The constructed presuppositions of the subjective will

Both the second and the third incarnations of the will (the will of an individual person and the will as an institutionalized, socio-ontological entity) are part of what Hegel calls the objective spirit and as a whole identifies with what he calls “the free will” (cf. Ency. §482).³⁵ Among other things, the domain of the objective spirit is the domain of the normative, which Hegel calls in its most abstract form “right” (*Recht*) (cf. Ency. §486). According to this, neither is normativity something over or above the objective, nor is the term “right” restricted to its juristic connotation. In a nutshell, “right” is the free will in an objectified form. The main idea of the concept of an objectified free will is the construction of a reality whose essential feature is that necessity (as an objective, formative principle) has to be understood as validity (*Gültigkeit*). Every entity that is part of this domain is already under some kind of implicit or explicit contract – sometimes but not necessarily in the form of a codified set of rules but at least in the form of a habitual mutual recognition of a certain set of rights and duties that constitutes and structures their community (cf. Ency. §§ 493–495). Hegel calls the first kind of will that falls under this domain the subjective (moral) will which he also denotes as the European concept of freedom (cf. Ency. § 503).

The internalization of good and evil is the central point of this kind of will. Evil is defined as the private, i.e. that which does not accord to the rules that are in force (i.e. the good). Good and evil are interdependent. Without the enactment of evil, every act would be either the satisfaction of a desire or an habitual fol-

³⁵ As mentioned in the first part of this paper: The second incarnation of the will (the will of an individual person) connects the anthropological and psychological foundations of the will (its first incarnation) with the socio-ontological realm Hegel denotes as the substantial will.

lowing of an implicit rule, a mere habit: Acting evil means acting against the valid, generally accepted right. In this sense acting evil means following a private rule according to a private interest that is only justified in respect to the subjective will of the person that acts upon it (cf. Ency. §§ 507–512).

The concept of “following a private rule” makes sense only insofar as it is negatively related to the right that it violates – outside of the boundary of this relation, a purely evil act would not be an act at all but just a mere behavior or perverted habit.³⁶ In this sense the genesis of an autonomous will that validates its own actions is modeled as the renunciation of the good as the (implicitly) generally accepted. Then again, this kind of evil is a presupposition for acting good: Only the real possibility to act evil makes it possible to act according to the good as the good, i.e. acting good according to one's own free will, to make the good one's own intention.

The subjective will is far from being a stable concept: To act is always to act with an intention. But by reflecting on the act, the individual subject will discover more than one intention that it (and others) can ascribe to itself, each of which is either good or bad depending on the context of other actions. Also the result of the action can differ from the intention up to the point where they contradict each other. From an extreme point of view, a concrete action is always bad (cf. Elements §§ 115–141).³⁷ This leads us to the third concept of will, the rational will as the constitution of a state in the realm of the ethical order (*Sittlichkeit*) (cf. Ency. §§ 513–552).

In the domain of morality the subjective will internalizes or incarnates the right (i.e. normativity) as the concept of good. The ethical order (*Sittlichkeit*) gives the different intentions of the subjective will a cohesive and consistent form and molds them into an objective reality which Hegel denotes as the actuality of freedom (cf. Ency. §§ 539,541). The beating heart of this reality is the constitutional state and the social entities it generates and those that stabilize it (family and civil society).

For Hegel, a legal constitution in its most general sense is not a contract between different persons but an authority that transcends the subjective intentions of those who recognize it (cf. Ency. § 540). It is (so to speak) the content of the good intentions of the subjective will that could not be realized by the

36 Hegel's remarks about the concept of evil revolve around Kant's conception of the pure evil in AA 6, 19–53, where he considers the extreme case of a will whose determination is completely unbound.

37 Robert Brandom's *A Spirit of Trust* (Brandom 2019) gives a detailed story of the instability of the content of the subjective will.

acts of a singular person. In this sense in the *Elements* Hegel calls the constitutional state the “substantial will” (*substanzielle Wille*) (Elements § 257).

Prima facie, this seems as if the constitutional state (and especially the highly problematic version Hegel outlines in the *Elements*)³⁸ could not be criticized by its subjects. But as we have seen, criticizing it is encoded in the very structure of the individual: The never ending transformation of social genetic presuppositions into constructed presuppositions is an integral part of the structure of the subjective will – a process that is an instantiation of the activity of self-preservation in the sense of the self-stabilization of the logical drive. I take this to be Hegel’s variation of Kant’s famous definition of enlightenment as “man’s emergence from his self-imposed immaturity” (Kant, AA VIII, 34). According to Kant, as its subjects we have to obey the regulations of the state, while as intellectuals we have the duty to criticize it. According to Hegel, the critique of our social reality is part of our very essence, of what it means to be an individual. The subjective will should in this sense be taken as the methodological starting point that we naturally inhabit when we think about our own condition. A standpoint that is ‘natural’ even though it is highly reflective.

Bibliography

- Brandom, Robert B. (2008): *Between Saying and Doing. Towards an Analytic Pragmatism*. Cambridge, Massachusetts: Harvard University Press.
- Brandom, Robert B. (2015): *From Empiricism to Expressivism. Brandom reads Sellars*. Cambridge, Massachusetts: Harvard University Press.
- Brandom, Robert B. (2019): *A Spirit of Trust: A Reading of Hegel’s Phenomenology*. Cambridge, Massachusetts: Harvard University Press.
- Chanter, Tina (2010): “Antigone’s Liminality: Hegel’s racial purification of tragedy and the naturalization of slavery.” In: Kimberly Hutchings/Tuija Pulkkinen (Eds.): *Hegel’s Philosophy and Feminist Thought: Beyond Antigone?* London: Palgrave-Macmillan, pp. 61–86.
- Churchland, Patricia S. (1987): “Epistemology in the Age of Neuroscience.” In: *The Journal of Philosophy* 84, No.10, pp. 544–553.
- Grabiner, Judith V. (1981): *The Origins of Cauchy’s Rigorous Calculus*. Cambridge, Massachusetts: MIT Press.
- Hegel, Georg Friedrich Wilhelm (1778): “Wissenschaft der Logik. Die Lehre vom Wesen (1813).” In: *Gesammelte Werke*. In Verbindung mit der Deutschen Forschungsgemeinschaft herausgegeben von der Rheinisch-Westfälischen Akademie der Wissenschaften. Vol. 11. Friedrich Hogemann/Walter Jaeschke (Eds.). Hamburg: Felix Meiner Verlag, pp. 233–409.

38 Cf. Poppers classic treatment in Popper 1945.

- Hegel, Georg Friedrich Wilhelm (1980): "Phänomenologie des Geistes." In: *Gesammelte Werke*. In Verbindung mit der Deutschen Forschungsgemeinschaft herausgegeben von der Rheinisch-Westfälischen Akademie der Wissenschaften. Vol. 9. Wolfgang Bonsiepen/Reinhard Heede (Eds.). Hamburg: Felix Meiner Verlag.
- Hegel, Georg Friedrich Wilhelm (1981): "Wissenschaft der Logik. Die Subjektive Logik oder die Lehre vom Begriff (1816)." In: *Gesammelte Werke*. In Verbindung mit der Deutschen Forschungsgemeinschaft herausgegeben von der Rheinisch-Westfälischen Akademie der Wissenschaften. Vol. 12. Friedrich Hogemann/Walter Jaeschke (Eds.). Hamburg: Felix Meiner Verlag, pp. 233–409.
- Hegel, Georg Friedrich Wilhelm (1984): "Wissenschaft der Logik. Die Lehre vom Sein (1832)." In: *Gesammelte Werke*. In Verbindung mit der Deutschen Forschungsgemeinschaft herausgegeben von der Rheinisch-Westfälischen Akademie der Wissenschaften. Vol. 21. Friedrich Hogemann/Walter Jaeschke (Eds.). Hamburg: Felix Meiner Verlag.
- Hegel, Georg Friedrich Wilhelm (1992): "Enzyklopädie der philosophischen Wissenschaften im Grundrisse (1830)." In: *Gesammelte Werke*. In Verbindung mit der Deutschen Forschungsgemeinschaft herausgegeben von der Rheinisch-Westfälischen Akademie der Wissenschaften. Vol. 20. Udo Rameil/Wolfgang Bonsiepen/Hans Christian Lucas (Eds.). Hamburg: Felix Meiner Verlag.
- Hegel, Georg Friedrich Wilhelm (2009): "Grundlinien der Philosophie des Rechts." In: *Gesammelte Werke*. In Verbindung mit der Deutschen Forschungsgemeinschaft herausgegeben von der Rheinisch-Westfälischen Akademie der Wissenschaften. Vol. 14,1. Klaus Grotzsch/Elisabeth Weisser-Lohmann (Eds.). Hamburg: Felix Meiner Verlag.
- Hegel, Georg Friedrich Wilhelm/Michelet, Karl L. (1986): "Vorlesungen über die Geschichte der Philosophie III." In: *Werke*. Auf Grundlage der Werke von 1832–1845 neu editierte Ausgabe. Eva Moldenhauer/Karl Markus Michel (Eds.). Frankfurt am Main: Suhrkamp.
- Hegel, Georg Friedrich Wilhelm (2010): *The Science of Logic*. George di Giovanni (Ed., Trans.). Cambridge, England: Cambridge University Press.
- Hegel, Georg Friedrich Wilhelm (2018): *The Phenomenology of Spirit*. Terry Pinkard (Ed., Trans.). Cambridge, England: Cambridge University Press.
- Heinrichs, Johannes (1974): *Die Logik der 'Phänomenologie des Geistes'*. Bonn: Bouvier Verlag.
- Henrich, Dieter (1967a): "Hegels Logik der Reflexion." In: *Hegel im Kontext*. Frankfurt am Main: Suhrkamp.
- Henrich, Dieter (1967b): "Hegels Theorie über den Zufall." In: *Hegel im Kontext*. Frankfurt am Main: Suhrkamp, pp. 157–186.
- Henrich, Dieter (1978): "Hegels Logik der Reflexion. Neue Fassung." In: *Die Wissenschaft der Logik und die Logik der Reflexion*. Bonn: Bouvier Verlag, pp. 203–324.
- Hoffmann, Thomas S. (1991): *Die absolute Form. Modalität, Individualität und das Prinzip der Philosophie nach Kant und Hegel*. Berlin, New York: De Gruyter.
- Houlgate, Stephen (1995): "Necessity and Contingency in Hegel's Science of Logic." In: *The Owl of Minerva* 27. No. 1, pp. 37–49.
- Husserl, Edmund (1982): *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy – First Book: General Introduction to a Pure Phenomenology*. F. Kersten (Ed., Trans.). Hague, Boston, Lancaster: Kluwer.

- Iber, Christian (1990): *Metaphysik absoluter Relation. Eine Studie zu den beiden ersten Kapiteln von Hegels Wesenslogik*. Berlin, New York: De Gruyter.
- Ihmig, Karl-Norbert (1989): *Hegels Deutung der Gravitation. Eine Studie zu Hegel und Newton*. Frankfurt am Main: Athenäum.
- Kain, Philip J. (2002): "Hegel, Antigone, and Women." In: *The Owl of Minerva* 33. No. 2, pp. 157–177.
- Kant, Immanuel (1907/14): "Die Religion innerhalb der Grenzen der blossen Vernunft (1793)." In: *Kants Werke*. Akademie-Textausgabe. Unaltered photocopy reprint of the text from the publication series of Kant's complete works initiated by the Prussian Academy of Sciences 1900 ff. Vol. 6. Preussische Akademie der Wissenschaften (Ed.). Berlin, New York: De Gruyter, pp. 1–202.
- Kant, Immanuel (1913): "Kritik der praktischen Vernunft (1788)." In: *Kants Werke*. Akademie-Textausgabe. Unaltered photocopy reprint of the text from the publication series of Kant's complete works initiated by the Prussian Academy of Sciences 1900 ff. Vol. 4. Preussische Akademie der Wissenschaften (Ed.). Berlin, New York: De Gruyter, pp. 1–163.
- Kant, Immanuel (1923): "Beantwortung der Frage: Was ist Aufklärung? (1784)." In: *Kants Werke*. Akademie-Textausgabe. Unaltered photocopy reprint of the text from the publication series of Kant's complete works initiated by the Prussian Academy of Sciences 1900 ff. Vol. 8. Preussische Akademie der Wissenschaften (Ed.). Berlin, New York: De Gruyter, pp. 33–42.
- Kornblith, Hilary (2002): *Knowledge and its Place in Nature*. Oxford: Oxford University Press.
- Kreis, Guido (2015): *Negative Dialektik des Unendlichen. Kant, Hegel, Kantor*. Berlin: Suhrkamp.
- Lampert, Jay (2005): "Hegel On Contingency, Or, Fluidity And Multiplicity." In: *Bulletin of the Hegel Society of Great Britain* 51, pp. 74–82.
- McDowell, John (1994): *Mind and World*. Cambridge: Harvard University Press.
- Merleau-Ponty, Maurice (2012): *Phenomenology of Perception*. Donald A. Landes (Ed., Trans.). London: Routledge.
- Mills, Patricia Jagentowicz (1986): "Hegel's Antigone." In: *The Owl of Minerva* 17. No. 2, pp. 131–152.
- Moretto, Antonio (1993): "Hegel on Greek Mathematics and the Modern Calculus." In: Michael J. Petry (Ed.): *Hegel and Newtonianism*. Dordrecht: Springer, pp. 149–165.
- Nuzzo, Angelica (2006): "Dialectic as logic of transformative processes." In: Katerina Deligiorgi (Ed.): *Hegel. New Directions*. Montreal: McGill-Queen's University Press.
- Paterson, Alan L. T. (1997): "Towards a Hegelian Philosophy of Mathematics." In: *Idealistic Studies* 27. No. 1/2, pp. 1–10.
- Popper, Karl R. (1945): *The Open Society and Its Enemies, Volume II: The High Tide of Prophecy: Hegel, Marx and the Aftermath*. London: Routledge.
- Quine, Willard V. (1969): "Epistemology Naturalized." In: *Ontological Relativity and Other Essays*. New York: Columbia University Press.
- Richli, Urs (1982): *Form und Inhalt in G. W. F. Hegels Wissenschaft der Logik*. München: Oldenbourg Verlag.
- Rosen, Stanley (2013): *The Idea of Hegel's Science of Logic*. Chicago: University of Chicago Press.

- Ryle, Gilbert (1938): "Categories." In: *Proceedings of the Aristotelian Society* 38. No. 1, pp. 189 – 206.
- Ryle, Gilbert (1953): *Dilemmas. The Tanner Lectures 1953*. Cambridge, England: Cambridge University Press.
- Searle, John R. (1983): *Intentionality. An essay in the philosophy of mind*. Cambridge, England: Cambridge University Press.
- Sellars, Wilfrid (1981): "Foundations for a Metaphysics of Pure Process – The Carus Lectures." In: *The Monist* 61. No. 1, pp. 3 – 90.
- Stekeler-Weithofer, Pirmin (1992): *Hegels Analytische Philosophie. Die Wissenschaft der Logik als eine kritische Theorie der Bedeutung*. Paderborn: Ferdinand Schöningh Verlag.
- Stekeler-Weithofer, Pirmin (2005): *Philosophie des Selbstbewusstseins. Hegels System als Formanalyse von Wissen und Autonomie*. Frankfurt am Main: Suhrkamp.
- Volkert, Klaus (1987): *Geschichte der Analysis*. Mannheim/Wien/Zürich: BI-Wissenschaftsverlag.
- Wandschneider, Dieter (1982): *Raum, Zeit, Relativität. Grundbestimmungen der Physik in der Perspektive der Hegelschen Naturphilosophie*. Frankfurt am Main: Klostermann.
- Wenz, Daniel (2016): "Wirklichkeit als Reflexionskategorie – Hegels normativer Realismus." In: Violetta L. Waibel/Max Brinnich/Christain Danz/et al. (Eds.): *Ausgehend von Kant. Wegmarken der Klassischen Deutschen Philosophie*. Würzburg: Ergon Verlag.
- Wolff, Michael (1986): "Hegel und Cauchy. Eine Untersuchung zur Philosophie und Geschichte der Mathematik." In: Rolf-Peter Horstmann/Michael J. Petry (Eds.): *Hegels Philosophie der Natur*. Stuttgart: Klett-Cotta, pp. 197 – 263.

Alex Englander

Hegel and the Paradox of *Willkür*

Abstract: Throughout his mature philosophy, Hegel deploys the terms *Willkür* and *willkürlich* with a markedly pejorative intent. In doing so, he has a variety of targets in his sights. When it comes to issues of philosophical methodology, for example, he frequently criticises certain ways of proceeding as *willkürlich* as a way of signalling that they fall short of his own conception of strictly philosophical science. And when it comes to issues in practical philosophy, he generally mentions *Willkür* to indicate an inadequate and misguided notion of free will. Without wanting to deny the overall consistency of Hegel's usage of these terms, I here restrict my focus to the latter issue. More specifically, in this essay I want to analyse Hegel's discussion of *Willkür* in his *Elements of the Philosophy of Right* in order to get a better handle on what he understands by the term. This will then allow us to explain one of the central claims of his discussion: namely, that if we understand the free will (*der freie Wille*) purely in terms of the freedom of *Willkür*, we will end up faced with a paradox. Unpacking this claim can, in turn, help us better understand Hegel's diagnosis of what he considers unsatisfactory accounts of free will and so, by way of contrast, put his own positive account in a clearer light.

1 Introduction

Our starting point is the fifteenth paragraph of the introduction to the *Philosophy of Right*, in which Hegel briefly weighs in on the “controversy” as to “whether the will (*Wille*) is actually free or whether our knowledge of its freedom is merely a delusion” (PR §15 A). At least in the days of 18th century rationalism, he thinks, this dispute understood the will (*Wille*) as *Willkür*. *Willkür* is often translated as “arbitrariness” or “power of choice”¹ and, as an initial approximation, we can

¹ The most recent translation of Hegel's Encyclopaedia *Philosophy of Spirit* (the W. Wallace/A.V. Miller translation, revised by M.J. Inwood and published as *Hegel's Philosophy of Mind*) translates the term as ‘wilfulness’. Insofar as this generally has pejorative connotations, suggesting that an agent is stubborn or headstrong, it is somewhat misleading. *Willkür* can of course bear this sense for Hegel, but only when a subject insists on their right to exercise their *Willkür* at the expense of the properly ‘ethical’ (*sittlich*). This attitude is therefore the practical counterpart, so to speak, of the conceptual error this paper addresses, namely, taking the freedom of *Willkür* to exhaust freedom as such. In this article, I simply stick to the German *Willkür*.

say that Hegel is addressing treatments of the will that place primary emphasis on the elective aspect of volition and so understand the will as an ability to decide voluntarily one way or the other, to make uncompelled choices. Hegel himself begins his discussion with a fairly casual designation of the *Willkür*, telling us that it is what people have in mind when they think of free will as “being able to do as one pleases”.

His judgment is that, in the first place, confusing freedom of the *Wille* with the freedom of *Willkür* betrays a “complete lack of cultivated thought” (*Bildung des Gedankens*). Be that as it may, Hegel tells us that free *Willkür* is the “commonest idea we have of freedom”, and we might well want to know whether our common understanding of free choice is real or delusory, however unsophisticated. Yet Hegel’s verdict on the controversy he mentions suggests that if we ask about the reality or unreality of free choice, we are already taking for granted that we have a coherent notion in view; but the notion of *Willkür* is in fact bound up with a certain *conceptual* problem:

To the certainty of this abstract self-determination, *determinism* rightly opposed the *content*, which, as something encountered, is not contained in that certainty and therefore *comes to it from outside* – although ‘outside’ here denotes drive or representation, or simply the fact that the consciousness is filled in such a way that its content is not derived from its own self-determining activity as such. Accordingly, since only the formal element of free self-determination is immanent within *Willkür*, whereas the other element is something given to it, *Willkür* may indeed be called a delusion if it is supposed to be equivalent to freedom (PR §15 A).

Much here requires elucidation, especially regarding the notion of ‘determinism’ at issue. Yet we can already note two points.

First, Hegel does not *directly* take a stand on the “controversy” – whatever precisely that controversy amounts to. That is, he does not claim outright that freedom of the will (*Wille*) *tout court* is a delusion and nor does he deny that there is such a thing as free *Willkür*. Rather, he reframes the issue in terms of a conditional claim: the idea of free *Willkür* is an illusion *if* it is taken to be definitional of free will as such. The implication is that we will arrive at a negative verdict on free will given a certain identification, namely of true freedom of the *Wille* with freedom of *Willkür*.

Second, at least the general shape of the resultant conceptual problem is clear, insofar as it revolves around a failure to integrate the *form* of elective willing, of voluntary choice, with the determinations of its content: the self-determination which characterises freedom of choice is undermined by the fact that its “content” is “given” to it “from outside”. And as external givens, they do not possess the self-given or self-“derived” character apparently needed for free will.

In what follows, I will elaborate Hegel's critique of the conflation of free will with free *Willkür* in the following steps: (2) will specify the conceptual problem at issue by expanding upon the threat presented by what Hegel here calls "determinism", showing it to have the form of a paradox; (3) will place the passage within the context of Hegel's broader discussion of *Willkür* in order to outline his fuller diagnosis of the relevant conceptual failure; (4) will then sketch certain features of Hegel's practical philosophy that are vital to transcending the understanding of free will in terms of *Willkür*. This final section will necessarily be somewhat programmatic, as its concerns nothing less than his general understanding of the relation between desires and their objects.

2 The Problem

It is important, first of all, to distinguish the threat posed to *Willkür* by "determinism" in §15 from that which governs most contemporary free will debates, namely, the threat of a specifically *causal* determinism. The latter threat stems from a collision between the commitments we have when we adopt respectively internal and external perspectives on our actions; that is, when we occupy the practical perspective *of* the deliberating agent versus when we adopt a theoretical, third-personal perspective *on* the agent's actions. The first perspective involves a commitment to our own freedom, arguably as a concomitant both of our first-personal stance as agents and of our attributions of responsibility. The second, by contrast, seems to demand a commitment to determinism in virtue of our belief in the explicability of natural phenomena.² The resulting conflict is a familiar enough starting point for many standard discussions of free will. The difficulty Hegel raises, by contrast, can get started without having to leave behind the internal, practical perspective of the deliberating agent; it can arise when an agent reflects upon the nature of their own motivational set. A certain course of reflection can lead to a paradox insofar as it sets down two apparently sound criteria of free agency, but draws an unacceptable conclusion on their basis.

² This paradox stemming from an internal-external conflict corresponds of course to several of Kant's presentations of the problem of free will (whatever his own ultimate stance on the status of this 'problem'), insofar as he is clear both that freedom is an indispensable condition of adopting a practical standpoint, and thus of willing, whilst thoroughgoing causal determinism is a necessary concomitant of our knowledge of appearances. See Kant 1996, 4:457 and 5:48f. *inter alia*.

In short, the problem is as follows: Thinking of freedom in terms of *Willkür* involves adopting the perspective of the acting subject and isolating a certain condition on something being a product of free agency; an agent knows herself to be free insofar as she is not *necessitated* to actualise any particular “content”, in Hegel’s terms, and so to take any particular course of action, but can *choose* to do as she wishes. What she does is free insofar as it is a result of her choice or decision. The object of the will here need not be a doing in the external world, at least not directly, as an agent can equally begin to reflect and look ‘inwards’ at the structure of her motivations; she can then decide that, say, some desire is or is not a suitable basis for action. Indeed, she can thus make first-order desires objects of free deliberation, by prioritising and organising her motivations so as to lend them a more ‘rational’ structure, by rendering them more consistent with her higher-order desires for example. Just like actions, grounds of actions do not have to be necessary givens for the subject, but can be subject to her choice.

We thus have two criteria of free volition in play. The first, which we can call the ‘mediation’ criterion, states that some event or state is a product of free agency if it is mediated by the agent’s choice; the second – call it the ‘grounding’ criterion – states that the agent make their choice on some basis (e.g. it does what it “pleases”). Yet these two seemingly non-negotiable criteria jointly imply that any resolution of the will must, at some point or other, depend upon given contents that are not a function of its choosing. Even if the basis of some choice is an upshot of a free reflective choice, there must at some point be a basis that is not such an upshot, but the brutally given basis of deliberation, something given but not subject to reflection. This conclusion is unacceptable because it undermines the mediation criterion, with the result that when an act is free insofar as it is a function of choice, it must be unfree insofar as the choice is unchosen. The essential idea here has been repeatedly thematised in the history of philosophy, receiving perhaps its most explicit formulation in Schopenhauer, who expands at length upon the claim that: “In a word, a human being always does only what he wills, and yet he necessarily does it. This is owing to the fact that he already is what he wills: for from what he is, all that he ever does follows of necessity” (Schopenhauer 1999, 88).³ The threat to freedom here, which clearly need not invoke causation at all, is what Hegel designates by “determinism” in §15.

3 “Mit Einem Wort: Der Mensch thut allezeit nur was er will, und tut es doch notwendig. Das liegt aber daran, dass er schon *ist* was er will: denn aus dem, was er *ist*, folgt notwendig Alles, was er jedesmal thut.” (Schopenhauer 1977, 138). For a variety of historical formulations of the paradox, see Strawson 2010, 41f.

In its more general form, the paradox plays upon the difficulty of finding a conception of the will that simultaneously satisfies two requirements, which reflect compatibilist and libertarian impulses respectively: on the one hand, since free actions are not mere reflexes or “mindlessly habitual” (Strawson 2010, 24), but intentional and performed for reasons, they must be intelligible in terms of an agent’s motivational set or character (in line with the grounding criterion); on the other hand, if actions are to be free, they seemingly cannot receive *full* explanations in terms of a prior motivational set or character, else the agent’s doings are ultimately the results of given factors (contrary to the mediation criterion). Along these lines, a more recent version of the paradox can be found in Strawson’s contention that true or ultimate self-determination is impossible. Beginning from the premises that “how one acts when one acts intentionally for a reason is, necessarily, a function of, or determined by, how one is, mentally speaking” and that responsibility for one’s actions – their being products of genuine self-determination – requires being responsible for “how one is, mentally speaking”, he derives a regress (Strawson 2010, 24–5). For responsibility for how one is entails having successfully brought about that state via a conscious choice (mediation); such a choice requires a principle of choice (grounding); and so responsibility for action requires having chosen this principle. Yet as soon as we apply the criterion of responsibility to someone’s being such a way that they choose this principle – namely, that they choose it in light of a further principle – the resultant regress is clear. Strawson is not especially specific about what constitutes a “principle of choice”, referring casually to “preferences, values, pro-attitudes, ideals, whatever” (Strawson 2010, 24), yet these correspond to what Hegel refers to as “content” in the passage cited above (and elsewhere calls “material” or “drives, desires and inclinations”, e.g. PR §11). The key point is that since some content must be given to the will qua content-choosing mechanism “from outside”, such that it chooses *on the basis* of something, it follows, in Strawson’s words, that “[w]e cannot possibly choose our root principles of choice, our conative base-structure, in the required way”, that is, to be truly, or ultimately, self-determining (Strawson 2010, 42).

This is the predicament the will faces if we conceive it purely as *Willkür*. Once we oppose the *Willkür* to a given “conative base-structure”⁴ and try to understand freedom in terms of the former’s elective capacity, the paradox necessarily

⁴ For the sake of maximal generality, I will from now on talk of ‘conative states’ or simply ‘desires’. My use of ‘desire’ is thus not restricted to the specific determination Hegel understands by that term in the *Encyclopaedia*, but simply picks out the various determinations of the practical will that *Willkür* sees as ‘content’.

ensues.⁵ On Hegel's view, it is only if we begin by modelling the will in terms of this opposition that we will find ourselves at the impasse just sketched. But, he thinks, there is nothing necessary about such a starting point when thinking about free will; it only seems to be given a certain interpretation of the phenomena of agency. Yet for Hegel's account of how we might arrive at this problematic interpretation, we must turn to the broader context of §15 and to his full discussion of *Willkür*.

3 The Derivation of *Willkür*

In the introduction to the *Philosophy of Right*, Hegel discusses *Willkür* as an initial, yet incomplete, realisation of the concept of free agency. He begins with a schematic description of free agency (which he calls “will”, *Wille*) in terms of what we might call generic reflective capacities. Firstly, one can abstract from any particular object of volition and think oneself purely as a universal ‘I’ (cf. PR §5). Thinking of oneself as such is to know that one is not compelled to actualise any object of volition. Secondly, a will must nevertheless determine itself to seek to realise one object or another – else it is not a will (cf. PR §6). Thirdly, Hegel claims that these two capacities are inseparable aspects of a unified phenomenon: minimally, I know myself as a universal I only insofar as I deliberately abstract from particular determinations; and I resolve upon and identify with a particular content in the knowledge that I, as a universal I, am not bound by any particularity (cf. PR §7).⁶

Even at this very early stage, before *Willkür* is introduced, we thus already have an indication of why it will exhibit a conceptually unstable structure. For Hegel's initial determination of the free will implies that free agency essentially involves adopting a self-conscious stance towards the phenomenon of agency. The will, for Hegel, is not any kind of substance; it does not have a nature that can be *independently* specified, be it in empirical or metaphysical terms, since an agent's perspective on their own agency is constitutive of agency itself.

5 For an alternative presentation of the paradox, associated with the “ultimacy clause” of an “ultimate responsibility” criterion of free will, see Kane 1998, 35–41. The ultimacy condition reads that “for every X and Y (where X and Y represent occurrences of events and/or states) if the agent is personally responsible for X, and if Y is an arche (or sufficient ground or cause or explanation) for X, then the agent must also be personally responsible for Y.” This of course leads to the introduction of an arche (Z) to make sense of responsibility for Y (ibid., 37).

6 For a fuller account of these opening paragraphs, see Ostritsch 2014, 55–69 and Yeomans 2012, 23–9.

In other words, as we will see in more detail, *what* the will is, is not something separate from how a subject conceives herself as a willing agent in adopting distinctively practical stances. The moment of abstraction described in §5, for example, is not a mere matter of the agent's stepping back from its needs, desires, drives etc., and surveying them as independent determinacies; it is not a matter of their looking at their 'will' from an external vantage point. Instead, Hegel describes it as an activity of dissolution (*Auflösung*), of deactivating its contents as immediate motivating grounds; it is an activity of the will itself, which is responsible for conative contents' having the very form they assume. What this form is we shall see shortly, but the immediate point is that the capacity of agency to abstract from its own contents, and to reflect on them as its contents, does not leave those contents unaffected; the will cannot adopt a position on itself without in the process shaping that on which it adopts a position.

Willkür is Hegel's name for the first such stance towards agency that he proceeds to discuss. It is the stance taken by the "reflective will", where the relevant kind of reflection is clearly that which he designates in the *Science of Logic* as "external reflection" (SL 348–55). The important feature of external reflection in the present context is that it precisely does purport to "stand over" its drives etc., as though conducting a neutral survey of self-standing, given items (PR §14), whose true nature is fully determinate and fixed *independently* of the dynamics of reflection itself. As such, it can reflect on its conative states, subject them to evaluative or critical consideration, and attempt to lend them a structure of its choosing, ranking and prioritising different motivations as it sees fit.

Willkür comprises two factors: "[a] free reflection which abstracts from everything, and [b] dependence on an inwardly or externally given content and material" (PR §15). As the stance sees free reflection confronted with different contents between which it can choose, it fits Hegel's schematic description of agency, knowing in its choosing that it is not inherently constrained to actualise any one specific content. Yet his description of the two factors is supposed to betray a failure of their integration: the stance oscillates between total abstraction from and dependence upon an arbitrary content, which "may or may not be mine" (PR §14). In the *Encyclopaedia*, Hegel expresses this oscillation as follows: "It is a contradiction: it actualizes itself in a particularity, which is at the same time a nullity for it, and has a satisfaction in the particularity which it has at the same time left behind" (Enc. §478). *Willkür* understands itself as utterly unrestricted, but has to succumb to dependence upon a content both circumstantially available and randomly chosen (cf. PR §16). From this stance, the will is a manifestation of contingency ("*contingency* in the shape of the will", PR §15), like a rolled dice, since any content is merely "a possible content in opposition to free

reflection" (PR §15 A) and is not a necessary content of the will qua abstract self-reflection.

This conception of agency as opposed to objects, whose contingency stems from their being "externally given", belongs to the reflective will's understanding of freedom as "formal self-activity" (PR §15 A). Hegel attributes the failure to integrate the two aspects of the free will to an overemphasis on its first aspect, abstraction, at the expense of its identification with contents. *Willkür* is accordingly a model in which all contents, all conative states, are external to agency, which thus shrinks down to a pure reflective capacity for electing independent contents. In introducing *Willkür* in §14, Hegel therefore says that it "*stands above* its content, i.e. its various drives, and also above the further individual ways in which these are actualized and satisfied." "Reflection", the self's conceiving of itself as merely "formal universality", thus introduces a double indeterminacy: it is utterly indifferent to the given natural, generic drives it takes itself to confront, and these are equally indifferent to any specific means for satisfying them. It is at this point that he introduces the claim that an identification of freedom with that of *Willkür* succumbs to a self-undermining paradox.⁷

Hegel is clear that the stance of *Willkür*, the particular interpretation of free will it represents, derives from presupposing a dichotomy of the universal 'I' and a manifold of *natural* givens: "The *reflective* will has two elements – the sensuous element and that of thinking universality". Given Hegel's definition of the free will, we are already led to expect that what the *Willkür* takes to be respectively internal (formal reflection) and external to the will (given sensuous content) is mediated result rather than immediate discovery; that what seems to be the agent's pre-constituted nature is in fact an artefact of its self-constitution.⁸ And he is therefore equally clear that this opposition is not simply *found* as the given constitution of agency, but instead "*produced by reflection*" (PR §21, my emphasis). It is thus a particular stance which generates a flat conative space, so to speak, in which objects are uniformly alien to free agency. Hence, as Christopher Yeomans points out, when Hegel shows how the stance of *Willkür* can be overcome, explaining how the drives "should be freed from the *form* of their immediate natural determinacy" (PR §19), he is highlighting how naturalness is a

⁷ In Wood's words, the problem of *Willkür* is that "[t]he object of an arbitrary choice has not been properly made part of the self that chooses it, since it bears only a contingent relation to the self's needs, desires and other choices" (Wood 1990, 49).

⁸ Hence, when deriving the reflective will in Enc. §474 Hegel says that the "inclinations and passions" merely "seem (*scheinen*) in their particularity to stand both to the individual and to each other in an external relationship and thus with an unfree necessity." Only as such are they "afflicted with contingency".

property of conative states' form, a *logical* feature, and not of their specific content or origin. The alien nature of the drive is not a result of its "provenance in the body in its causal history", but of how *Willkür* opposes it as a contingent particular to the universal simplicity of its elective activity (cf. Yeomans 2012, 28).

It is important to note that Hegel explicitly associates the standpoint, and indeed paradox, of *Willkür*, with the "philosophy of reflection" and, as ever, mentions Kant as a philosopher who remains at this standpoint and shares its formalistic understanding of freedom (cf. PR §15 A). Associating Kant's view of freedom with arbitrariness may seem odd, not only because Kant sees self-subjection to the moral law as constitutive of freedom, but also because Hegel relies heavily on Kant's practical philosophy throughout the opening of the *Philosophy of Right*. Yet, regardless of the accuracy of Hegel's Kant interpretation, we can make sense of this reference by reminding ourselves of Kant's famous definition of autonomy from the *Groundwork*: "the property the will has of being a law to itself (independently of every property belonging to the objects of volition)" (Kant, 4:440).⁹ Hegel, as a good post-Kantian, recognises the central importance of the initial part of this definition, insisting that the will can have nothing but its own freedom as its object (cf. PR §27), but rejects the Kantian construal of the phrase in parentheses. In "objects of volition", "objects" refers to the desired states of affairs, meaning that "volition" here is willing based upon inclination. When an agent is moved by an inclination to realize some state of affairs, their maxims, in Henry Allison's words, "reflect the agent's needs as a sensuous being, needs that are themselves explicable in terms of the laws of nature" (Allison 1990, 97).

My first-order desires, those that make certain courses of action salient as options, – i.e. provide me with what Kant labels "incentives" – are grounded in my needs as a natural being, and since they are apt for an explanation based upon laws of nature, insofar as my maxims reflect them alone, the law is given *to* me, not self-given. But this means that Kant's understanding of conative states operates on the assumptions of *Willkür*: we are to understand these states by abstracting from the agent's distinctively rational, reflective capacities, and, to maintain their naturalistic explicability, regard them as essentially independent of their objects. We may of course train our sensuous desires so that we are inclined to do what the law requires, but the relation between both desires *in themselves* and their objects, and between conative contents quite generally and

⁹ All references to Kant are to the volume and page of *Kants gesammelte Schriften*, herausgegeben von der Deutschen (formerly Königlichen Preussischen) Akademie der Wissenschaften, 29 vols. (Berlin: Walter de Gruyter, 1902).

rational reflection, is arbitrary and contingent. And it is only against this set of background presuppositions, Hegel believes, that it becomes necessary to seek an a priori norm internal to practical reflection just as such. Given this background, the nature of our 'given' desires is then as relevant to our autonomy as any other natural state or process, such as digestion, so that what really matters is not *which* desires we have, but the *manner* in which we integrate them into formal principles constitutive of the reflective capacity alone and subordinate certain kinds of incentives to others.¹⁰

Hegel's own means of avoiding the paradox is to reject this background picture. That is, he does not think that unconstrained rational reflection suspended between worldly objects on the one side and a "conative base-structure" on the other represents the original or inherent nature of agency. And if we need not see the will as inherently separate from, and trapped in a hopeless endeavour to catch up with a given base-structure, then nor need we worry about facing the paradox head-on. In his view, reflection in the guise of *Willkür* takes a *particular stance* towards agency and reifies it into the putatively given structure of practical subjectivity as such. More specifically, as we will see in more detail below, he believes that reflective thought has a tendency to pick out one – admittedly vital – practical capacity amongst others that rational subjects develop (namely *Willkür* qua capacity for deliberative, critical reflection), and absolutises it, thus conflating free *Willkür* with free *Wille* as such. Further, Hegel thinks that we develop this particular capacity against the background of more fundamental stances, stances that involve a quite different relation between rational reflection, desires and their objects. We can now turn to fleshing out certain elements of this alternative picture.

10 Kant's first systematic use of *Willkür* as a technical term occurs in the *Religionsschrift*, where he invokes a power of choice in order to explain how we both "adopt" a highest "evil maxim" and overcome our tendency [*Hang*] to evil by adopting a "good disposition". In this explanation, the *Willkür's* ability to form and select maxims that subordinate different kinds of incentive, such that one conditions the other, is central. "Depravity", for example, is thus "the propensity of the power of choice to maxims that subordinate the incentives of the moral law to others" (Kant, 6:30; see too 6:37). Strawson (2011, 42) believes that Kant's insistence that moral character (*Gesinnung*) is ultimately self-incurred shows him biting the bullet by appealing to a paradoxical self-creation: "The human being must make or have made himself into whatever he is or should become in a moral sense, good or evil. These two [characters] must be an effect of his free power of choice (*Willkür*), for otherwise they could not be imputed to him and, consequently, he could be neither morally good nor evil" (Kant, 6:44).

4 Diagnosis

Hegel's discussion of *Willkür* soon gives way to the demand that "the drives should become the rational system of the will's determination" (PR §19) in a way that shows the *Willkür* to have mistaken its own nature. Trapped in its paradox, *Willkür* at first tries to strip its contents of their arbitrariness and contingency (to "purify" them, PR §19) by organising them in accord with an abstract aim that is of its own choosing: happiness (*Glückseligkeit*) (PR §20). But this ploy soon turns out to be self-defeating, as either the aim is so abstract as to be uninformative, or it borrows concrete content from some particular given material which, since it too will be a contingent given, just takes *Willkür* back to square one (PR §20 A). Through the failure of the concept of happiness to serve as a criterion of choice, it becomes apparent that *Willkür* was never really aiming simply to fulfil its first-order desires, or to achieve "happiness" as their abstract totality, but to maintaining itself as a structure of self-determination through different contexts; what matters for the will that conceives itself as *Willkür* is not primarily doing as it pleases but, so to speak, *being able* to do as it pleases.¹¹ Its true object, in other words, is its own freedom (PR §21; 27). Ultimately, the conclusion Hegel is working towards is that the will can have freedom as its object only within the institutions of ethical life in which, he says, the opposition of naturalistically explicable givens and abstract reflection will be overcome and revealed as an abstraction from a more complex unity: "This universality [in which the will is its own object] is such that the *immediacy* of the natural and the *particularity* with which the natural is likewise invested when it is produced by reflection are superseded (*aufgehoben*) within it" (PR §21).

How this 'supersession' occurs is, of course, a longer story. The crucial point for our purposes is that the step beyond *Willkür* that sets the stage for the rest of the *Philosophy of Right*, the demand that the drives become "the rational system of the will's determination", is vulnerable to a certain misinterpretation. We will not supersede the standpoint of *Willkür*, that is, if we interpret the demand as meaning that the rational will remains external to the drives and submits them to its own rank-ordering or organising principles, so that the subject comes to find herself inclined to do what rationality dictates *anyhow*. In that case, far from comprising a rational system, the drives would remain blind, merely borrowing the reflective will's quite independent rationality. On the metaphysics of *Willkür*, it is indeed the case that drives and desires, as natural givens, have no constitutive relation to their objects. As we have seen, within the

11 For a helpful presentation of this progression see Houlgate 1995, 869.

interpretive stance of *Willkür*, the drives' relation to objects is just as contingent as their relation to reflection; different objects simply may or may not happen to satisfy them. Hegel's way forward is not to argue that the previously empty reflective somehow comes to find self-standing, substantial principles in the institutions of ethical life, and that it becomes 'rational' by moulding its desires into a motivational set that remains inherently blind, but inclines the subject in the direction of ethical behaviours. On the contrary, put highly schematically, his argument is, firstly, that the "reflective will" only emerges from *within* contexts of ethical life in the first place; it is not as though a fully functioning practical reason could predate ethical enculturation and then *discover* substantial principles by landing on the right institutions; and secondly, practical reasoning within ethical contexts fundamentally involves a responsiveness to the way in which conative states present objects and situations in inherently evaluative terms. To remain at this schematic level, practical reasoning is then not a matter of manipulating conative states from without, of compelling them along channels moulded by a self-sufficient capacity for practical reason, but of responding appropriately to the directives of the drives as an *already* rational system.¹²

Without being able to go into Hegel's theory of *Sittlichkeit* here in any detail, we can make the second of these two ideas somewhat clearer by considering how *Willkür* has to understand the relation between the indeterminate desires it posits and the specific objects to which a subject comes to establish a relation of desiring: however it explains this connection, it has to conceive the resulting desire as a kind of compound. Since a conative state is essentially indeterminate with respect to its objects, it has an entirely generic character, which allows of individuating descriptions independently of the particular objects that might satisfy it (hence they can enter into naturalistic explanations). And this is part of why *Willkür* is "*contingency* in the shape of the will": it is not only that whichever

¹² McDowell articulates a version of the present point in his interpretation of Aristotle's claim that the "desiderative element" (the *orektikon*) can be persuaded by practical wisdom. On McDowell's interpretation, this means that the "desiderative element" is the "seat of directive rationality", rather than just "obediently receptive" to it. McDowell wants to argue that the intellectual states of practical wisdom could not have their content without a habituation of the *orektikon* (McDowell 2009, 50f.). The parallel with Hegel is of course unsurprising given the Aristotelian concern with second nature he shares with McDowell. In his brief discussion of virtue at the beginning of the *Philosophy of Right*'s section on *Sittlichkeit*, Hegel stresses the indispensable conative element of ethical life, explaining how the "subject bears spiritual witness to them [the "ethical powers"] as to its own essence, in which it has its *Selbstgefühl*" (PR §147). Via habituation, the ethical "appears as *custom*; and the *habit* of the ethical appears as a *second nature* which takes the place of the original and purely natural will and is the all-pervading soul, significance, and actuality of individual existence" (PR §147).

desire “formal self-activity” happens to land upon is ultimately a contingent matter, but that that desire in itself has no necessary relation to its object. But with *Willkür*’s conception of the relation in question, the very arbitrariness of the relation between desire and object means that the prioritisation of any objects of desire over any others must itself be arbitrary (just as we find when *Willkür* invokes the notion of happiness).

That contradiction which is the *Willkür* (see §15) makes its appearance as a dialectic of drives and inclinations which conflict with each other in such a way that the satisfaction of one demands that the satisfaction of the other be subordinated or sacrificed, and so on; and since a drive is merely the simple direction of its own determinacy and therefore has no yardstick within itself, this determination that it should be subordinated or sacrificed is the contingent decision of *Willkür* (PR §17).

But by precluding our appreciation of how conative states can indeed present a “yardstick” (*Maß*) or standard, the stance of *Willkür* falsifies the nature of desires by means of a generalisation that prescind from their object-directedness. We can motivate this point by building upon some considerations of Matthew Boyle. Boyle’s target is Christine Korsgaard’s theory of autonomy. – He worries that her manner of conceiving the relationship between desire and rational reflection involves a similar lack of integration. The picture he wishes to dispel is one in which we possess a power, essentially of a kind shared with other animals, to be impelled by brute desires, whilst reason functions as an additional power to monitor and regulate that independent conative system. This picture would be incompatible with Korsgaard’s talk of desires “bidding” for and *with-standing* reflective endorsement.¹³ In Hegel’s terms: how can something stand up to rational scrutiny if it is a blind drive, “merely the simple direction of its own determinacy”? We must therefore, Boyle recommends, conceive desires as presenting their objects as *prima facie* “to-be-pursued” and thus as meriting desire (Boyle 2014, 451). They present themselves as meeting standards independently of the rational reflection which might then subsequently endorse or reject that putative claim to suitability.

Boyle cites Quinn’s notorious example of the man neuro-physiologically (re) constituted so as compulsively to turn on any radio he sees. Whereas Quinn’s aim

¹³ For Korsgaard’s talk of desires bidding for our endorsement, see e.g. Korsgaard 1996, 94–7. Given Hegel’s account of *Willkür* as a mode of external reflection, it is revealing how Korsgaard describes the stance of practical reasoning in a way that evokes the image of standing over one’s conative states: “When you deliberate, it is as if there were something over and above all of your desires, something which is *you*, and which *chooses* which desire to act on” (Korsgaard 1996, 101).

is to refute the Humean thesis that a brute desire could function as a reason for action, Boyle denies that such an impulse could even be a “normal human desire” (Boyle 2014, 451). The “mode of presentation” of the radio compulsive’s desires should be modelled, he says, on the way nausea presents a need to vomit, that is, as “tending to become unavoidable” (Boyle 2014, 450). The respective impulses provide causal explanations of why some behaviour may occur, but they do not present them as meriting some performance. I may entertain some additional reason to indulge or permit the impulse, but insofar as it is not a normal desire, I am more likely to have a reason to take some means to cut off the impulse at the source.

We might want to elaborate Boyle’s account by drawing a further distinction between two senses of the conative mode of presentation of “to-be-pursuedness”. Both contrast with cases in which reflection encounters an operative need it would prefer to dispel, rather than something claiming worthiness to be desired. Firstly, there is the kind of physiological compulsion which serves for Boyle’s own contrast cases. Someone might think: ‘I wouldn’t now keep hiccupping if I didn’t feel so compelled to and knew that resisting it would just cause more trouble etc.’ This thought, however, displays a different logical form to a different case, for example: ‘I wouldn’t be doing this if I didn’t really need the money.’ Unlike the hiccupping example, this relation to an experienced need is not strictly parallel to the nausea case. The need for money is no mere *cause* of the pursuit. If it presents the pursuit as unavoidable, this is in a different sense to the felt inevitability of hiccupping. The motivation for the pursuit may well present itself as worthy of endorsement, in light of its efficacy, the agent’s situation, and so on. Yet the person would still not bother to engage in the pursuit if the money she needed suddenly fell from the heavens; the pursuit is only desirable in light of its happening to be the most efficacious, convenient etc. means of satisfying some anterior desire.

The reason for making this further distinction is that it allows us to bring more clearly into view the kind of desire Boyle likely has in mind, but which surely does not exhaust the class of “normal” desires. The person who reluctantly pursues some end because of the money does not, unlike the radio compulsive, fall outside the domain of the normal. Yet their case contrasts with a logically prior class of cases in which desires present their objects as worthy of pursuit, and thus of meriting desire, *in their own right*. This is the way in which they present agents with self-sufficient reasons for action, which, at least insofar they present genuine reasons, agents are entitled to act upon without further reflection. Call such desires “primary” desires. The kind of relation of desire to object they embody clearly differs both from that in which the desirability of some course of action, say, is purely parasitic on an anterior aim, and

from that in which the ‘desire’ is better classified as a compulsion or impulsive need, whose relation to its object can be exhaustively specified in purely causal terms. In both of these common enough cases, the relation between desire and object is external, and the desire does not require any particular satisfaction, remaining essentially indifferent to its object. It may be that there is, as a matter of contingent fact, only one way of satisfying some need (of dispelling nausea, for example), but this is very different from the way in which primary desires present the inherent meritoriousness of their objects. In the latter case, the relevant mode of presentation is one where the to-be-pursuedness of *this particular* object is the very way in which the agent possesses the conative state itself. With primary desires in the picture, therefore, we can remove *Willkür*’s understanding of the object-directedness of conative states, on which that relation is at bottom a kind of contingent compound. Accepting the results of Boyle’s line of reasoning, we can allow for conative states to be inherently relational and object-directed, both necessarily related to their objects and essentially individuated by them.¹⁴

On Hegel’s view, it is within the institutions of ethical life that subjects acquire the conative profiles that allow for an identification with actions on the basis that they are the realisation of desires that present their objects as worthy of pursuit in their own right. An ethical subject, for whom performing such actions has become “second nature”, and so standardly performs them without necessarily engaging in explicit reflection, still meets Hegel’s minimal criteria of free agency from §5–7 of the *Philosophy of Right*: they know in their acting, at least implicitly, that the content they are actualising represents a possible content amongst others insofar as, whilst identifying with the content, they are at the same time a universal “I”, and thus not wholly absorbed within it. Now, it is because this minimal structure of free agency remains satisfied that the (for Hegel, ethically vital) capacity for deliberative reflection has space to intervene and to interrupt the subject’s ongoing acquiescence with what their desires at least seem to present as meritorious.¹⁵ In such reflection, a subject can question

¹⁴ Hegel therefore both maintains that a range of conative states – from sensations, to feelings, to passions etc. – are individuated by their objects and relate to these as form to matter, and insists repeatedly that these states become the forms of, for example, ethical and religious contents. For a discussion of how the dialectic of subjective spirit exhibits this complex form-content structure, see Rometsch 2007, 74 f.; and for how ‘lower’ determinations can be the form of ethical and religious contents, see Halbig 2002, chs. 2–3, especially his discussion of how Hegel would reject any strong thesis of the mind’s modularity.

¹⁵ Of course, Hegel thinks that the conditions on free agency are in fact only truly fulfilled within ethical life. It is not as though inhabiting ethical institutions is merely compatible with the

whether or not some object does indeed merit desire and can come to a negative conclusion. Upon critical examination, a subject might diagnose a desire as a mere compulsion or need, as something that seemed to reflect the rationality of the will, but in fact does not. From this standpoint, the “drives, desires and inclinations” can take on the guise of being merely “sensuous”, “natural”, and “particular” contents, of being independent magnitudes with no intrinsic connection to the rational will. In other words, they can assume the *logical* form that is bestowed upon them by the interpretive framework of *Willkür*. Reflection is therefore in part a realisation of agency’s fundamental abstractive power mentioned in §5: it is a power to alienate the will from its own contents, to suspend their normative force, and to consider them as the contingent causal grounds of our choices and actions, which may well explain, but do not of themselves justify them.

At this stage, it is important to distinguish the power of reflection to alienate contents, which Hegel does see as the exercise of *Willkür*, from the position he attacks. The latter is what I have been referring to as the standpoint or perspective of *Willkür*, which is assumed by the “philosophy of reflection” and conflates free will as such with the “free reflection” that *Willkür* represents. Whilst in one respect, as we shall see, marking any strict separation here is somewhat artificial, we should consider *Willkür* as a capacity and *Willkür* as a theoretical standpoint in turn.

We noted above how Hegel thinks of free agency in terms of self-conceptions of free agency, as stances that the will adopts toward itself. As a capacity to alienate and reflect critically upon conative states, *Willkür*, as we have seen, involves conceiving the will as “formal self-activity” or “free reflection” and conceiving its possible contents as generic, particular givens. These two elements are in fact aspects of one and the same achievement: extracting the will as a pure faculty of reflective choice from its immersion in ethical life is accomplished in and through stripping desires of their status as “measures” or “yardsticks” and seeing them as simply a “sensuous element”. In his explanation of the development of *Willkür* in the *Encyclopaedia*, Hegel explains how the will “distinguishes itself from the particularity of the urges, and places itself as simple subjectivity of thinking above their diversified content” (Enc. §476). And as he stresses both in the *Encyclopaedia* and in the *Philosophy of Right*, the “diversified content” assumes the form of contingent “particularity” in and through this act

concept of free will as he begins to unfold it in the introduction to PR. Rather, ethical life represents the actualisation of the concept of free will, forming the basis of agents’ ability to rationally identify with the contents of their wills as at the same time a merely possible content.

of the will's conceiving itself as "simple subjectivity"; attaining this practical self-conception as "free reflection" goes hand in hand with deactivating the normative character of desires, giving them the logical form of the natural. It is through this process, then, that practical *Geist* can become "the standpoint of choosing between inclinations" (Enc. §476), the standpoint from which the paradox of *Willkür* can get a purchase.

I remarked earlier that there is something superficial about insisting on too stark a contrast between *Willkür* as a capacity and *Willkür* as a theoretical standpoint. This is because adopting a certain standpoint is clearly integral to the former. The critical reflection that *Willkür* enables requires an agent to see that reflection as independent of and set over against a series of optional natural contents between which it can choose. The error of *Willkür* as a theoretical standpoint is, as stressed above, the specifically philosophical manoeuvre of interpreting this setup as the fundamental, given nature of agency itself. Of fundamental importance to Hegel's critique of *Willkür* as a theoretical standpoint is thus that it ignores how *Willkür* as a mature capacity of deliberative reflection is something subjects have to *cultivate*, through practice, against the background of ongoing ethical commitments, of a practical orientation that requires a basis of what we have called primary desires, however developed. *Willkür* as a theoretical standpoint is an outgrowth of the reflection constitutive of *Willkür* as a capacity when it goes so far as to reify its operative picture, as though the will just is, essentially and originally, "formal reflection" suspended between conative givens on the one hand and given contingent objects on the other. Part of this reification is to take the content or "material" of the will, the "drives, desires and inclinations" as they appear to *Willkür*, and to hypostasise them into a "conative base-structure".

Hegel is clear that the ability to subject ongoing commitments to reflective critique and to diagnose them as ultimately contingent elements of our motivational set is indispensable from an ethical viewpoint¹⁶; but this is no reason to think that the capacity to place conative states in reflective suspension exhausts free will or that free action can only be action that flows from the position from which such standpoint-free reflection is conducted.¹⁷ Doing so presupposes a

¹⁶ It is particularly important for developing the rights of "subjectivity" of the specifically *moral* agent. See esp. PR §§121–24.

¹⁷ Robert Pippin thus puts the danger of taking an "abstractly free will" in isolation as follows: "The way in which various inclinations and aversions are put temporarily out of play, while manifesting what Hegel refers to as an 'infinity,' the term he often uses for a complete self-authorization and so self-determination, gives us only a partial picture because such a deliberative suspension is itself always a manifestation of a finite position as well, always institutionally and

metaphysical picture of the will that is based on a generalisation of an alienated position, one which prescind from how our conative states present us with reasons in cases of ordinary ethical commitment. It supposes that the will can have a free relation to its “material” only if the latter is an object of choice – a supposition that invites the regresses discussed in (2). Yet as inhabitants of ethical institutions, our being responsive to reasons is a matter of our responsiveness to how conative states present their objects as meriting our engagement – and indeed in their own right, not primarily as satisfying brutally given conative states to which they are indifferent. As Hegel puts it in the *Encyclopaedia*:

Their genuine rationality [viz. of drive and inclination, A.E.] cannot emerge in an examination conducted by *external* reflexion, which presupposes *independent* natural determinations and immediate drive, and therefore lacks the single principle and final purpose for them. But it is the immanent reflexion of the mind itself, going beyond their *particularity*, beyond their natural *immediacy*, and giving their content rationality and objectivity, in which they become *necessary* relationships, rights and duties. (Enc. §474 A, translation amended).

From the standpoint of external reflection, both the capacity of *Willkür* and a conative base-structure seem to be “immediate” ingredients of subjectivity in whose constitution reflection plays no role. Yet since *Willkür* is a cultivated capacity, one that is cultivated on the basis of an abstraction of “formal reflection” from an immersion in normatively rich conative states, any such base-structure is itself an abstraction. Positing it is to ignore what the *Willkür* as an abstractive, deliberative capacity is always abstracting from viz. the ongoing commitments of ethical life. *Sittlichkeit* provides the necessary background without which the *Willkür* would have nothing to lend the form of mere givenness, and so it is a mistake to think of the contents of volition as a natural base-structure to which a “formal” capacity of reflective choice can be subsequently added. Hence, as we saw in section (2), Hegel insists that the ‘naturalness’ of the objects of *Willkür* is a function of reflection, a feature of their logical form, not a matter of their ontological or genetic status.¹⁸ Conflating the two is to invite the paradox of *Willkür* and bring free will into question.

historically bound, never as abstract and complete as this first picture presents it.” (Pippin 2008, 37).

18 Yeomans draws attention to an important dimension of *Sittlichkeit*’s being the background of our exercise of *Willkür* (albeit in other terms) with his Hegelian counterpart to Kant’s ‘incorporation thesis’: the “expulsion thesis” explains how it is through ethical action and commitment that agents make and remake the distinction between what counts as a rational, normative content and what as merely sensuous. According to this thesis, “the distinction between reason and

Of course, I have said nothing concrete about ethical life, about the “*necessary* relationships, rights and duties” that form the content of the truly free will in Hegel’s view. An examination of his positive theory of freedom would have to consider his account of these institutional arrangements and the theory of rationality that undergirds it. But I have aimed to provide something of a *via negativa* to such an examination by outlining Hegel’s critique of an alternative way of thinking about the freedom of the will – a way that is perhaps even more widespread in our time than in his own.

Bibliography and Abbreviations

- PR: Hegel, Georg Wilhelm Friedrich (1991): *Elements of the Philosophy of Right*. Allen W. Wood (Ed.)/H.B. Nisbet (Tr.). Cambridge: Cambridge University Press. [Cited by section number.]
- Enc: Hegel, Georg Wilhelm Friedrich (2007): *Hegel’s Philosophy of mind*. W. Wallace/A.V. Miller (tr.), revised with an introduction and commentary by M.J. Inwood. Oxford: Clarendon Press. [Cited by section number.]
- SL: Hegel, Georg Wilhelm Friedrich (2010): *The Science of Logic*. George di Giovanni (tr.). Cambridge: Cambridge University Press. [Cited by page number.]
- Allison, Henry (1990): *Kant’s Theory of Freedom*. Cambridge: Cambridge University Press.
- Boyle, Matthew (2014): “Additive Theories of Rationality: A Critique”. In: *European Journal of Philosophy* 24, pp. 527–555.
- Halbig, Christoph (2002): *Objektives Denken: Erkenntnistheorie und Philosophy of Mind in Hegels System*. Stuttgart-Bad Cannstatt: Frommann-Holzboog.
- Houlgate, Stephen (1995): “The Unity of Theoretical and Practical Spirit in Hegel’s Concept of Freedom”. In: *The Review of Metaphysics* 48, pp. 859–881.
- Kane, Robert (1996): *The Significance of Free Will*. New York; Oxford: Oxford University Press.
- Kant, Immanuel (1996): *Practical Philosophy*. Mary J. Gregor (Ed./Tr.). Cambridge: Cambridge University Press.
- Kant, Immanuel (1996b): *Religion and Rational Theology*. Allen W. Wood/George di Giovanni (Ed./Tr.). Cambridge: Cambridge University Press.
- Korsgaard, Christine (1996): *The Sources of Normativity*. Cambridge: Cambridge University Press.
- McDowell, John (2009): *The Engaged Intellect: Philosophical Essays*. Cambridge, Mass.: Harvard University Press.
- Ostritsch, Sebastian (2014): *Hegels Rechtsphilosophie als Metaethik*. Münster: Mentis.

nature within the moral psychology of the agent is something created by rational action, rather than a presupposed field of representational inputs of differing source and value”. See Yeomans 2015, 29–35.

- Pippin, Robert B. (2008): *Hegel's Practical Philosophy: Rational Agency as Ethical Life*. Cambridge: Cambridge University Press.
- Rometsch, Jens (2007): *Hegels Theorie des erkennenden Subjekts: systematische Untersuchungen zur enzyklopädischen Philosophie des subjektiven Geistes*. Würzburg: Königshausen & Neumann.
- Schopenhauer, Arthur (1976): *Zürcher Ausgabe: Die beiden Grundprobleme der Ethik: behandelt in zwei akademischen Preisschriften. Band 6*. Angelika Hübscher (Ed.). Zürich: Diogenes.
- Schopenhauer, Arthur (1999): *Prize Essay on the Freedom of the Will*. Eric F.J. Payne (tr.)/Günter Zöller (Eds.). Cambridge: Cambridge University Press.
- Strawson, Galen (2010): *Freedom and Belief*. Oxford: Oxford University Press.
- Yeomans, Christopher (2012): *Freedom and Reflection: Hegel and the Logic of Agency*. Oxford: Oxford University Press.
- Yeomans, Christopher (2015): *The Expansion of Autonomy: Hegel's Pluralistic Philosophy of Action*. Oxford: Oxford University Press.

Jenny Bunker

Ethics and Will in Schopenhauer's Philosophy

Abstract: Schopenhauer explains compassion as the expression of the profound understanding which recognises that individuals are phenomenal: at the level of the will reality is undivided. This explanation undermines the distinction between egoism and compassion. It is argued that a broader conception of the ethical allows Schopenhauer's account of compassion to be reconciled with his metaphysics. His theory of salvation, too, can only be understood in the light of his philosophy of will, but these elements of his thought also come into conflict. Salvation is supposed to involve bliss and the abolition of the will – this is explicable only on a modest reading of the will's metaphysical status according to which it is neither the thing-in-itself nor foreign to individuation. The paper concludes with a methodological suggestion on how best to approach the contradictions between Schopenhauer's metaphysics and his ethics.

1 Introduction

For Schopenhauer, ethical value attaches only to compassion and to asceticism. He is at pains to point out the coincidence of this aspect of his moral philosophy with a variety of ethical and religious systems which similarly hold these traits in high regard. Schopenhauer's own accounts of compassion and asceticism, though, are entirely conditioned by his metaphysics of will. That is, Schopenhauer's description of the nature of the world as will determines his pictures of the ethical modes of compassion and asceticism. At the same time, a number of paradoxes arise with his theories both of compassion and of salvation through the ascetic denial of the will which seem to indicate that they are inconsistent with the very metaphysical system in which they are grounded. I focus on three in particular. First, the clear moral distinction Schopenhauer wishes to draw between compassion and egoism is undermined by his own explanation of what makes compassion possible. Second, the promise of a blissful state of salvation cannot be fulfilled in a world that comprises the striving and suffering will. Third, Schopenhauerian salvation depends on the abolition of the will, but the world will has not been (and will not be) abolished.

In the case of compassion, I will argue, the best that can be done in Schopenhauer's defence is to overhaul his account of compassion's value and of the

scope of the ethical as such. The case of asceticism and salvation is still more interesting and significant: taking his picture of them seriously pushes us towards a particular understanding of the nature of the Schopenhauerian will and its role in his philosophy.

I will start by indicating three of the most salient features of Schopenhauer's account of the will. I will then set out the way in which his model of compassion is determined by his theory of will, how, precisely, the theory of will undercuts the model of compassion and the best prospect for a solution to this. Next, I will briefly outline Schopenhauer's account of the possibility of salvation through the ascetic denial of the will. I will detail two paradoxes that arise in relation to this: that concerning how and when the Schopenhauerian subject can experience bliss, and that concerning the possibility of the will's abolition. I will suggest that the best prospect for resolving these paradoxes is to modify our understanding of the role of the will in Schopenhauer's philosophy. I will develop this putative solution a little further in exploring links with and between other commentators' readings of Schopenhauer's account of salvation. I will finish with a comment on different strategies for defending Schopenhauer's ethics of will from charges of paradox, indicating my own preferred approach.

2 The Schopenhauerian Will

Of particular relevance to this discussion are the following features of the Schopenhauerian account of will. First, for Schopenhauer, the inner nature of every individual and the inner nature of reality itself is will. Indeed, in his great leap beyond Kant, Schopenhauer claims that the will is the thing-in-itself, for instance in the *World as Will and Representation* (Volume I): “the will is *thing-in-itself* [...] the innermost essence, the kernel, of every particular thing and also of the whole” (Schopenhauer 1969 [1818/19], 110). Through the knowing subject's imposition of the forms of the principle of sufficient reason, the will expresses itself in the individuated representations of ordinary experience: these constitute the world as representation. Second, the forms imposed by the subject include space and time. Jointly, these two compose the principle of individuation. This means that the will itself, which is not subject to these forms, is beyond plurality and differentiation. Third, the will's fundamental character, and that of all willing beings, is unending striving, which in turn means that suffering is ubiquitous and unavoidable. To strive is to lack, to lack is to suffer: each of us is an instantiation of the striving will, each of us is fated to meaningless and unending suffering. This is the centrepiece of Schopenhauer's storied pessimism.

The will, then, can be identified as the thing-in-itself: it constitutes the inner nature of the world and everything in it. The world as will is beyond plurality, but thanks to the organising activity of knowing subjects, in particular their imposition of the forms of space and time, the will appears as distinct individuals in the world as representation. The will's essence is striving, hence its individual expressions strive and suffer.

3 Compassion

3.1 The Paradox of Schopenhauerian Compassion

Nonetheless, despite his pessimism, Schopenhauer identifies two genuinely valuable modes of being for individual subjects: compassion and asceticism. He introduces compassion as, alongside egoism and malice, one of the three basic character-types, and the only one which is morally commendable. Schopenhauer is perhaps best described as a kind of virtue-ethicist: certainly, he thinks the consequences of an action are irrelevant to its moral value (and nor does duty play any fundamental role). Moral value, for him, attaches only to the character-driven desire or intention which drives an action, and the only morally worthy source of motivation is compassion. Therefore, Schopenhauer names compassion the sole mark of morality.

It is difficult to see that he gives us adequate grounds to accept that compassion is the *sole* mark of morality and I am inclined to agree with Michael Maidan's assessment that this assertion of Schopenhauer's is "plainly erroneous." (Maidan 1988, 266). I will return to this point at the end of my discussion of Schopenhauer's ethics of compassion, but will focus for now on a different problem: that of how compassion can be possible *at all* if reality is as Schopenhauer describes. There is good reason to believe that his theory of the will destabilises his ethics of compassion, rendering the latter inexplicable by centring the role of egoism in human behaviour. For an action to be done out of compassion is, for Schopenhauer, for it to be done from a concern for the well-being or suffering – the weal or woe, as he puts it – of another. At the same time, in *On the Basis of Morality* he describes *egoism* as "[t]he chief and fundamental incentive in man" and says that "[a]s a rule [...] all man's actions spring from egoism" (Schopenhauer 1998, 131). Of course, there may be a degree of Schopenhauerian pessimistic hyperbole here, but egoism is at the very least the easier incentive to explain, given his metaphysical voluntarism.

This is a point that has troubled many of Schopenhauer's interpreters. Maidan asks "[i]f egoism is a general characteristic of being, how is altruism possible

at all? It would seem, at first glance, that altruism is something *contra naturam*, a mere impossibility” (Maidan 1988, 268). Christopher Janaway notes that “[c]ompassion therefore seems to go against our nature as individuals” (Janaway 2009, 7) and David Cartwright remarks that, given his metaphysical monism, “it appears that Schopenhauer should have been a motivational monist in some sense of that term, recognizing that there was some single motive that ultimately accounts for all human actions” (Cartwright 2009, 139). If all reality is ultimately an expression of striving will (this is Schopenhauer’s metaphysical monism) it might seem to follow that all human behaviour would be driven by egoistic striving. If so, the motivational pluralism which attributes some actions to altruistic compassion stands in need of explanation.

To reiterate: our fundamental nature is to be creatures of will, striving for our own ends. This is because reality itself is will, characterised by an endless striving, and expressing itself as striving individuals. Every individual feels the identity of his own will with *the* will. Moreover, the world as representation exists only for and in total dependency upon the representing subject, so that each individual also knows himself to be, as representing subject, the supporter of the entire phenomenal world. Therefore, Schopenhauer himself explains, “every individual [...] makes himself the centre of the world, and considers his own existence and well-being before everything else”. This rises to the point, Schopenhauer says, of an individual countenancing the annihilation of the entire world to maintain his own existence a little longer (Schopenhauer 1969, §61).

What, then, leads this arch-pessimist to conclude that altruism is real, and that there are acts which are genuinely undertaken out of compassion for others? The answer seems in part to be empirical: he simply enumerates certain behaviours which must have been driven by compassion, because they can be explained in no other way. In *On the Basis of Morality* he calls these “isolated but indubitable cases”, each as rare as a four-leafed clover (Schopenhauer 1998, 126). Of course, these appeals to experience might not be convincing, and readers may well feel that such instances as a poor man returning a rich man’s lost property could indeed be explained as springing from egoistic motives. Accordingly, David Cartwright notes that in *On the Basis of Morality* Schopenhauer admits his examples of altruism are “not immune to doubt” and that others may be able to explain them egoistically, but refuses to debate with such people, limiting himself to addressing “those who admit the reality of the matter” (Cartwright 2009, 140 quoting Schopenhauer BM, 139/203). This perhaps hints that he has separate reasons for believing in the reality of compassion as a motive force. Certainly, Schopenhauer needs to do more than simply *state* that compassionate incentives exist – it behoves him to *explain* how that is possible for striving, will-driven creatures.

He does this by an appeal to the more profound and truer metaphysical knowledge possessed by the compassionate person as compared to the egoist. Ordinarily, we experience reality as divided up into distinct individuals. This is because space and time – jointly forming the principle of individuation – are forms of experience applied by the knowing subject. The egoist is completely bound up in this world-view. He construes the world as essentially divided into individuals, naturally focuses his attention on the individual which is himself and is constantly aware of what threatens it. Because of this, as Schopenhauer puts it, the egoist “feels himself surrounded by strange and hostile phenomena, and all his hope rests on his own well-being” (Schopenhauer 1969, §66). But Schopenhauer’s own metaphysical position identifies phenomenal individuals as illusions that deceive those who see the world through the veil of Maya. In reality, everything is will, to which the principle of individuation does not apply. Compassion and goodness are, to Schopenhauer, the product of this more profound knowledge, which sees through phenomenal individuation to our unity in the will. In §64, Schopenhauer writes of a “deeper knowledge, no longer involved in the *principium individuationis*, a knowledge from which all virtue and nobleness of mind proceed” and in §66, he insists that “virtue must spring from the intuitive knowledge that recognizes in another’s individuality the same inner nature as one’s own”. Indeed, Schopenhauer says, “To be cured of this delusion and deception of Maya and to do good works of love are one and the same thing” (Schopenhauer 1969, §66).

According to Schopenhauer, I am *really* as much you as I am me and to hurt another is to hurt myself: “tormentor and tormented are in themselves one [...] This will likewise suffers both in the oppressed and in the oppressor” (Schopenhauer 1969, §64). The upshot of this, though, is that compassion appears to collapse into a kind of enlightened self-interest. This seems to be confirmed by Schopenhauer’s suggestion that the compassionate person is “just as little able to let others starve, while he himself has enough and to spare, as anyone would one day be on short commons, in order on the following day to have more than he can enjoy.” (Schopenhauer 1969, §66) Given the understanding that we are all bound up together at the level of the will, the compassionate person’s concern for another becomes a version of the prudent person’s self-interested action. Arguably, then, Schopenhauer ought to hold not that a compassionate person is someone of a radically different character to an egoistic person, but that compassion results from the combination of a self-interested nature and knowledge of how the world really is.¹

1 David Cartwright notes that “[t]his type of analysis” – that is, that compassion involves pen-

If compassion is to remain distinct from egoism then, as critics including Chris Janaway, David Cartwright and Bernard Reginster have noted, the compassionate person must see the object of her concern as a separate person.² If compassion is to be explicable, given the ubiquity of egoism, it can only be because the compassionate person sees no such distinction. We seem to be left either with accepting the existence of compassionately-motivated behaviour as a blunt and inexplicable fact (supported by Schopenhauer's disputed empirical examples) or with an explanation of compassion which makes it ultimately impossible to distinguish clearly from egoism. No wonder he calls it 'the great mystery of ethics'.

3.2 Broadening the Conception of Ethics

A conflict emerges, as we have seen, between Schopenhauer's fidelity to altruistic compassion as the *sine qua non* of morality and his own metaphysics of the will, according to which each of us strives constantly for what we desire. The combination of his metaphysics and epistemology *does* allow Schopenhauer to give a genuine explanation of how, in some, this could entail striving for the well-being of "others". Some individuals understand reality in a more profound way than most of us: they penetrate the veil of Maya and recognise our unity at the level of the will. While a typical individual identifies overwhelmingly with her phenomenal self, truer knowledge leads a person to identify with and so strive towards the interests of all. But what this picture undercuts is Schopenhauer's right to insist on the inherently *moral* chasm between compassionate and egoistic motivations. They are indeed distinct epistemologically and also practically (in terms of the actions which flow from each), but the compassionate person appears just as much as the egoist to be motivated by the pursuit of her own

etrating the *Principium Individuationis* and seeing another's woes as my own – "has led commentators, such as Eduard von Hartmann, Friedrich Nietzsche, Max Scheler, Patrick Gardiner, and D. W. Hamlyn to suggest that Schopenhauer reduced compassion to some form of egoism" (Cartwright 2009, 144). This is also an argument that Julian Young makes, claiming that despite Schopenhauer's official position "the altruist *does* act for the sake of his own interest, the only difference between him and the egoist being that he acts for the sake of the interests of his *metaphysical* rather than his empirical self" (Young 2005, 182).

2 So for instance Reginster notes that "Schopenhauer himself insists, in my view correctly if perhaps not consistently, that compassion requires a clear consciousness of the difference between us and the other: 'at every moment, we remain clearly conscious that *he* is the sufferer, not *we*; and it is precisely in *his* person, not in ours, that we feel the suffering [...]' (*BM* 147)" (Reginster 2009, 116 footnote 1).

well-being, the real difference is that she has a truer understanding of the nature of her self.

The problem for Schopenhauer, therefore, is that while he holds that the only good action is that motivated by compassion, his account of the compassionate person is of one who sees through the illusory distinction between self and other, realises our suffering is intertwined and so seeks the well-being of all. The distinction between the egoist and the compassionate person, then, seems to be an epistemological rather than a moral one. Ultimately, it looks as though all actions must be self-seeking: the problem, for Schopenhauer, is his insistence that such actions cannot be morally good.³

It is extremely difficult to find any reason given in the *World as Will and Representation* to support Schopenhauer's insistence on compassion as the sole mark of morality. In fact, it seems to run against his own statement that "every good is essentially relative; for it has its essential nature only in its relation to a desiring will" (Schopenhauer 1969, 362). If good and bad are defined in relation to the will and there can be no standard independent of the interests of individuals, it is unclear what standard Schopenhauer can use against which to measure compassion as morally good and egoism as bad (or at best non-moral).

Schopenhauer is often keen to draw connections between his own moral philosophy and other ethical and religious traditions. In this case, as Michael Maidan puts it, "there is a contradiction between Schopenhauer's account of morality and his description of the history of ethical thought." In particular, "Schopenhauer rightly points out that classical ethical thought was eudemonistic [...] pity entered into Western ethical thought only through Christianity" (Maidan 1988, 266). Schopenhauer's assertion that nothing other than pity or compassion can have moral value betrays an understanding of the ethical that is narrow to the point of idiosyncrasy. Furthermore, there are, of course, well-established ethical systems, including those modelled on flourishing, for which acting in one's own self-interest is compatible with – even necessary for – virtue. Figures including Seneca, Aristotle, Spinoza, Mill and Nietzsche held self-interest to be in harmony with morality.

Schopenhauer gives us no convincing reason to accept his picture of moral value as exhausted by compassion and antithetical to egoism. If he would countenance a wider understanding of ethics, he could explain both how compassion differs from egoism (in its care for an expansive, non-individuated self, recognised through a more profound metaphysical understanding) and the possibility

³ Saying of a morality that motivates "by acting on self-love" that "what springs from this has no moral worth" (Schopenhauer 1969, 367).

of its existence against the backdrop of the world as will. As it is, Schopenhauer holds that nothing has moral value that fails to spring from a compassionate will. I have suggested that he gives no satisfying reason for this stipulation, and indeed that the profound distinction he draws between compassion and egoism is undermined by his own metaphysics of will as well as his claim that good is relative to desire.

4 Salvation

4.1 Salvation and the Will

Schopenhauer's ethical theory – and indeed, the *World as Will and Representation* itself – culminates in a portrait of salvation through ascetic denial of the will. Despite his profound pessimism, Schopenhauer completes his philosophical system with the promise that, to a few, real, unalloyed bliss is available. Schopenhauerian salvation works as follows: an individual penetrates the veil of Maya to a more complete extent than the compassionate person. This occurs either through profound knowledge or extreme personal suffering. She understands the illusory nature of individuation, the underlying metaphysical unity of will and the inevitability of suffering. This picture of horror and futility causes the person to withdraw: her knowledge becomes a “quieter of the will” (Schopenhauer 1969, 392). She comes to deny the will through ascetic renunciation of pleasure and refusal of what is necessary to sustain existence. Ultimately, she starves to death.

Clearly, this model of salvation only makes sense in the context of Schopenhauer's philosophy of will. Salvation is *necessary* because reality is suffering, because of the striving nature of will. Ascetic withdrawal is *prompted* by a dawning awareness of this reality. And it *consists* precisely in denial of will – of a willing being renouncing willing itself.

Several questions immediately arise: I will focus on two. First, why is this process something to be desired – something deserving of the name salvation? Second, how is the death of the ascetic to be distinguished from that of the suicide, which Schopenhauer condemns as futile? The answers to these questions highlight notorious paradoxes with Schopenhauer's account of salvation. My argument will be that the best hope of resolving these depends on a different and more modest understanding of the role of will in Schopenhauer's metaphysics. In particular, I will suggest that salvation could only be understood as blissful on the assumption that will is not the ultimate reality, and that if the ascetic's

death is marked out by achieving the abolition of the will, then the Schopenhauerian is obliged to countenance individuation at the level of the will.

4.2 The Bliss of Salvation

In terms of the first question, Schopenhauer describes salvation as a blissful and enviable state that “we cannot behold without the greatest longing” (Schopenhauer 1969, §68). He speaks not only of an escape from suffering but of a life he characterises as “full of inner cheerfulness and true heavenly peace”. Indeed in the same section, he talks of bliss and blessedness, inferring from the “most blissful” (*säligsten*) moments we experience – in aesthetic contemplation – to just “how blessed [*wie sälig*] must be the life of a man whose will is silenced [...] forever” (Schopenhauer 1969, §68). Thus it seems for Schopenhauer that some of us can hope for a longed-for state of cheerfulness, peace and, indeed, blessedness.

When, though, is this bliss supposed to be achieved – during the life of the ascetic saint or after death? There appear to be problems with either option. Renunciation of the will, Schopenhauer tells us, consists in a constant struggle against new desires, in voluntary starvation and in the seeking out of pain. It is hard to see how this could constitute a blissful condition. On the other hand, if Schopenhauerian salvation is pictured as a post-death experience, *this* is very difficult to explain on his terms. It is correct to say that, for Schopenhauer, one's true self survives death, but by one's true self he means will. The Schopenhauerian self exists as knowing subject or intellect and as willing subject. The intellect is the servant of the will and does not survive the body's death. In Volume II of the *World as Will and Representation* Schopenhauer makes his departure from predecessors clear on this point: “[a]ll philosophers have made the mistake of placing that which is metaphysical, indestructible, and eternal in man in the *intellect*. It lies exclusively in the *will*” (Schopenhauer 1966, 495).⁴ To repeat, then, the knowing subject does not survive death. One's inner being is not intellect but will, and it is this will which is untouched by the death of the phenomenal individual. But if it is only as will that “we” survive, that is really just to say that the will is timeless and unaffected by the passing away of one of its phenomena. Admittedly each of us *is* will, will is our inner

⁴ Knowledge, consciousness and the intellect end with the brain's death: “[c]onsciousness [...] consists in knowledge; but this, as has been sufficiently demonstrated, belongs, as activity of the brain, and consequently as function of the organism, to the mere phenomenon, and therefore ends therewith” (Schopenhauer 1966, 495).

being, but as Dale Jacquette stresses, nothing that is unique to each of us survives: “[t]he part of me that survives death [...] is no different from but exactly the same as the pure willing part of you that survives death” (Jacquette 2005, 124–5). There is nothing that is unique to the individual which survives nor, it seems, any kind of subject who could experience bliss.

4.3 The Abolition of the Will

And in fact, in the case of the ascetic, death appears to be even more absolutely an obliteration than in any other. This takes us to the second question. Schopenhauer says of suicide that it is futile, since it destroys merely the phenomenon and not the inner nature: “the thing-in-itself remains unaffected” (Schopenhauer 1969, 399). By contrast, in the case of the ascetic’s death, he says, a “complete abolition” of the will is achieved (Schopenhauer 1969, 412). In this category alone “[i]t is not merely the phenomenon, as in the case of others, that comes to an end with death, but the inner being itself that is abolished” (Schopenhauer 1969, 382). In this one case, will as well as phenomenon is destroyed with death.

The difficulty here concerns what Schopenhauer means by this last claim – whether it is the *ascetic’s* will, or the will as such that is abolished. The will is supposed to be one, not divisible into parts. If the ascetic truly destroys *his* will, he then surely destroys *the* will as such – and with it the whole phenomenal world which is its manifestation. Indeed Schopenhauer insists that “the whole phenomenon of the will [...] the universal forms of this phenomenon [...] all these are abolished with the will. No will: no representation, no world” (Schopenhauer 1969, 411). His remark that “the rest of nature has to expect its salvation from man who is at the same time priest and sacrifice” does seem to confirm that thanks to the ascetic who attains to denial the abolition of the will and its phenomena is achieved, releasing them from their suffering (Schopenhauer 1969, 381).

If this reading is correct, however, then the defender of Schopenhauer needs to explain the continued existence of the world, since Schopenhauer appears to believe that there have been ascetic saints who have achieved full denial and the abolition of the will (see in particular Schopenhauer 1969, 401–402). Further, if anyone is ever to attain this state even in the future, it is presumably right to say that on Schopenhauer’s account the world couldn’t currently exist. He tells us past and future come into being with the first representing subject, they must surely also disappear with the abolition of the will. Given the current existence of phenomenal reality, the conclusion must be that no-one has attained and no-one ever will attain Schopenhauerian salvation. The alternative to this is to ac-

cept that it is the ascetic's own will, not the world-will as such, which is abolished with the completion of ascetic renunciation.

4.4 Resolving the Paradoxes of Salvation

On this latter account, then, the denial and abolition of the ascetic's will need not – indeed, does not – destroy the will as such. One might just possibly infer from Schopenhauer's comment that “for him who ends thus, the world has at the same time ended” that what he had in mind was that the successful ascetic destroys *his* will, and the world is abolished *for him* (Schopenhauer 1969, 382). This alternative may seem more appealing for another reason: Schopenhauer's account of moral responsibility, appearing earlier in book four of the *World as Will and Representation*, also seems to rely on individual wills being separable from the will as such. This is because he holds that each of us has moral responsibility for the actions which spring from his or her character, because a person's character is a free act of will. My actions within the world as representation are determined, but this does not immunize me from responsibility because they are determined in part by my own character, which is itself undetermined. Each person's character is a unique, individuating act of will. Given there are different characters, it appears there must be differentiation at the level of will itself.

For Schopenhauer, to reiterate, there can only be moral responsibility where there is freedom, but everything within the world as representation is determined, so only the will is free. Each of us is morally responsible because each has – more correctly *is* – a free intelligible character. Schopenhauer asserts that “the intelligible character of every man is to be regarded as an act of will outside time” (Schopenhauer 1969, 289). This is free because it is an act of *will* (falling outside the forms of the principle of sufficient reason, only *within which* laws of causality apply) but note, too, that it is *an act* of will. In other words, my intelligible character, my will, is an individual. It is also, perhaps, worth noting that Andrew King adduces a third reason to countenance individuation at the level of the will. If a subject can temporarily cease to will in aesthetic experience, he argues, it must be possible to distinguish modes of will (King 2005, 259).

The best prospect of resolving the second of these paradoxes, then, is to accept that the will *is* divided into individual acts constituting distinct intelligible characters. It is the ascetic's own will which is abolished in death: this one case remains qualitatively different from any other death in which only the phenomenon is destroyed, but it does not entail the destruction of the world-will as such. So the most fruitful reading of Schopenhauer's account of salvation requires us

to reenvision the will as not existing entirely beyond the principle of individuation. In other words, to bring it a little closer to the phenomenal and a little further from the thing-in-itself.

Returning to the first paradox of salvation – concerning how there can be the potential for bliss in a reality whose essence is the eternally striving will – two comparatively recent commentators have suggested a similar strategy. Schopenhauer says that in the experience of salvation we “see the world melt away with the abolished will, and retain before us only empty nothingness” (Schopenhauer 1969, 411). This is where the experience of bliss comes in, of “deep tranquillity [...] and serenity”. For Julian Young, this requires that will is not, in fact, the ultimate reality. Rather, that, as Schopenhauer puts it, what is left when the will is abolished is only nothingness relative to our own worldly perspective. Young reminds us that, in Volume II of *WWR*, Schopenhauer retreats from his claim that the thing-in-itself is will and accepts that the thing-in-itself is unknowable, leaving room for the possibility that ultimate reality is not the evil of the irrational and striving will, and that a blissful experience of reality is at least possible (Young 2009, 164). Similarly, as Andrew King parses it, the will is not (as Schopenhauer sometimes mistakenly says) the thing-in-itself-as-it-is-in-itself, but only the thing-in-itself in relation to appearance. For Schopenhauer, just as for Kant, the thing-in-itself as such is unknowable and, writes King, “not a legitimate object for the discipline of philosophy” (King 2005, 257). Philosophy can say nothing about the thing-in-itself-as-it-is-in-itself – we can only conceive it as the negation of the phenomenal, as a relative nothing, so it remains possible that there are modes of existence unknowable and incomprehensible to us (King cites *World as Will and Representation* Vol. II, 196–197 in support).

For salvation to be possible, according to the proposed response to these two paradoxes, the will must be subject to division, so that the ascetic saint’s will can be abolished, and it must not be the ultimate reality, so that there could potentially be some divine one-ness behind it. It is only by taking a more modest reading of the will that we can hope to show Schopenhauer’s metaphysics to be compatible with his doctrine of salvation. Given that the world exists, the will as such has not been abolished. The abolition achieved by the ascetic saint, then, must be the abolition of his own will. This means that the will cannot be entirely foreign to individuation. Given that salvation involves bliss, but no bliss is possible within the confines of the striving will, there must be a reality beyond the will: the will cannot be the thing-in-itself.

This approach, then, offers a solution to two of the major paradoxes associated with Schopenhauer’s account of salvation. First, it renders the continued existence of the world compatible with Schopenhauer’s view that there are people who have achieved salvation – salvation involves the abolition of an individ-

ual will, not the world will as such. Nonetheless, it retains the idea that the ascetic's death is qualitatively different from any other: it comprises the abolition of will, not simply of phenomenon. Second, this approach offers a way to make sense of the possibility of bliss given the evil and suffering nature of the will. Were the will to exhaust reality, no bliss would be possible; if there remains a thing-in-itself beyond or behind the will, that possibility is revived. Each of these putative solutions depends on a similar strategy: a deflationary reading of the metaphysical status of the will, deposing it from its position as thing-in-itself, thereby cohering with Schopenhauer's discussion of relative nothingness at the end of WWR I and with what he (disingenuously) claims in WWR II had always been his position on the status of the will.

4.5 A Final Mystery: The Subject of Salvation

However, there remains an outstanding problem concerning the subject who could experience this unknown reality or relative nothing. The death of the ascetic, as I understand it, sees the destruction of both the knowing and willing self. Possibly, if this reality in-itself is beyond the will, Schopenhauer envisages some noumenal subject who could experience it as bliss. This is, of course, a possibility about which we can say very little, assuming we can make any sense of it at all. The only Schopenhauerian subjects of whom we have knowledge are the knowing subject and the willing subject. If there is a noumenal realm beyond the will, it must surely be, as it is for Kant, utterly unknowable for us.

We have touched already on Andrew King's view that talk of the thing-in-itself exceeds the remit of philosophy – his own suggestion is that Schopenhauer's thought at this point becomes mystical rather than philosophical. This chimes with G. Steven Neeley's prior suggestion that we should understand Schopenhauer as speaking in a non-philosophical register in his account of the ascetic's deliverance from suffering. Where King compares Schopenhauer to the Christian mystic Meister Eckhart, Neeley argues that understanding it in the light of the profound influence Eastern thought had on his work allows for a non-paradoxical reading of Schopenhauer's account of salvation. "The Eastern religions maintain that the true self – as identified with the ultimate reality – is beyond the reach of all assertive statements," he writes (Neeley 1994, 118). Similarly, Schopenhauer's account of the death of the ascetic should, according to Neeley, be understood metaphorically – a litany of commentators have accused Schopenhauer of paradox and contradiction because they have taken him to be speaking literally. Moreover, Buddhist doctrine recognises a true self which

it identifies with “the Supreme Reality which lies beyond the phenomenal world” (Neeley 1994, 124). Schopenhauer might understand the subject who experiences salvation in this way, or, more likely – according to Neeley – insist that the “entity which survives natural death is beyond our comprehension” (Neeley 1994, 125) and leave “the mystic, not the philosopher, to proceed positively where knowledge proper has reached its limit” (Neeley 1994, 126).

These interpretations allow Schopenhauer to escape the charge of contradiction in his picture of a blissful experience of salvation, but at the expense of the prospect of giving any *philosophical* reason to believe in the possibility of such an experience. An alternative response would be that it makes most sense to think of the *knowing subject* as being the subject who experiences the bliss of salvation. It is the knowing subject who experiences the temporary happiness of an escape from the pain of willing during aesthetic experience, an experience which Schopenhauer alludes to as giving a foretaste of salvation. As noted above, he remarks that from the bliss of aesthetic experience “we can infer how blessed must be the life of a man whose will is silenced” and is, through asceticism, “left only as pure knowing being” (Schopenhauer 1969, §68). The price of construing the subject of salvation’s bliss as the knowing subject, however, is that it cannot be permanent but must end with death. Death obliterates the knowing subject, so this “most blissful” state could only be experienced by the knowing subject during the period at the end of life when the ascetic’s will has been silenced “except for the last glimmering spark that maintains the body and is extinguished with it” (Schopenhauer 1969, §68, 390). At death, the ascetic, uniquely, achieves complete annihilation – both as knowing and as willing subject. That in itself is an advantage over those of us doomed to suffer after death in new phenomenal forms – whether the ascetic also has the hope of a blissful post-death experience is not, it would seem, a question that philosophy can answer.

5 Conclusion

I conclude with a methodological note on interpretations and defences of Schopenhauer’s philosophy. In both the case of compassion and that of salvation, we have an ethical concept which emerges out of and can only be explained against the backdrop of Schopenhauer’s philosophy of will, but whose relation to the will is paradoxical. In the first case, Schopenhauer’s explanation of compassion depends on his picture of metaphysical unity in the will, but runs up against his insistence on the striving nature of the will and the ubiquity of egoism. In the second, salvation consists in denial of the will and is needful because of its

evil nature, but this very fact seems to render the blissful character of salvation inexplicable. Given the remarked clashes between his ethical doctrines and his metaphysics, there seem to be three broad strategies available to someone who wishes to defend Schopenhauer. The first is to abandon the relevant ethical position, the account of compassion or of asceticism respectively, to save the philosophical system. The problem with this strategy, of course, is that Schopenhauer insisted upon the systematic nature of his philosophy – even describing the *World as Will and Representation* as the outworking of a single thought. Moreover, the ethical doctrines are supposed to form the apex of that system. Not only would hiving off compassion and salvation leave Schopenhauer with an unremittingly bleak picture, it would be to excise what he himself names the “most serious” part of his work (Schopenhauer 1969, 271).

The second possible strategy is to see what can be salvaged of the philosophy of value if we assume the metaphysics fails – an approach which seems to have found favour recently. It is inspired by the idea that Schopenhauer overreaches in his metaphysical leap beyond Kant but that there is much that is worthwhile in his aesthetics and ethics, so if these come into conflict with his metaphysics it is the latter that should be sacrificed. Certainly, Schopenhauer would object to the idea of taking the ethical doctrines apart from his system as such and in this paper, I have tried to show that, in fact, the Schopenhauerian concepts of compassion and salvation only have meaning in relation to his philosophy of will.

The third option is to see if one or other – the ethical concept, or the underlying metaphysics – can be reworked to make the two compatible. This is the strategy I have tried to pursue. In the case of compassion, I would argue that the best prospect is to refigure the notion of compassion by giving up the insistence on its absolute distinction from egoism. This allows for compassion to be explicable on Schopenhauerian terms: the compassionate person feels with others and seeks their well-being because he or she recognises our unity in the will and so no longer draws a clear distinction between self and other. In the case of ascetic salvation, my suggestion is that the most promising approach is to take a more modest view of the role of the will in Schopenhauer's philosophy, accepting that it cannot be the ultimate reality, the thing-in-itself, nor entirely foreign to the forms of the principle of sufficient reason.

Bibliography

- Cartwright, David (2009): "Compassion and Solidarity with Sufferers: The Metaphysics of Mitleid". In: Neill, Alex/Janaway, Christopher (Eds.): *Better Consciousness: Schopenhauer's Philosophy of Value*. Chichester: Wiley Blackwell, pp. 138–156.
- Jacquette, Dale (2005): *The Philosophy of Schopenhauer*. Chesham: Acumen.
- Janaway, Christopher (2009): "Schopenhauer's Philosophy of Value": In: Neill, Alex/Janaway, Christopher (Eds.): *Better Consciousness: Schopenhauer's Philosophy of Value*. Chichester: Wiley Blackwell, pp. 1–10.
- King, Andrew (2005): "Philosophy and salvation: The Apophatic in the Thought of Arthur Schopenhauer". In: *Modern Theology* 21:2, pp. 253–274.
- Maidan, Michael (1988): "Schopenhauer on Altruism and Morality". In: *Schopenhauer Jahrbuch* 69, pp. 265–272.
- Neeley, G. Steven (1994): "A Critical Note on Schopenhauer's Concept of Human Salvation". In: *Schopenhauer Jahrbuch*, pp. 97–128.
- Reginster, Bernard (2009): "Knowledge and Selflessness". In: Neill, Alex/Janaway, Christopher (Eds.): *Better Consciousness: Schopenhauer's Philosophy of Value*. Chichester: Wiley Blackwell, pp. 98–119.
- Schopenhauer, Arthur (1966): *The World as Will and Representation Volume II*. Trans. Eric F.J. Payne. Toronto: Dover.
- Schopenhauer, Arthur (1969): *The World as Will and Representation Volume I*. Trans. Eric F.J. Payne. Toronto: Dover.
- Schopenhauer, Arthur (1998): *On the Basis of Morality*. Trans. Eric F. J. Payne. Indianapolis: Hackett Publishing.
- Young, Julian (2009): "Schopenhauer, Nietzsche, Death and Salvation". In Neill, Alex/Janaway, Christopher (Eds.): *Better Consciousness: Schopenhauer's Philosophy of Value*. Chichester: Wiley Blackwell, pp. 157–170.

Index of Names

Allison, Henry 8, 40–42, 49, 51, 65, 67,
99, 182–183, 235

Blumenbach, Johann Friedrich 160

Bouterwek, Friedrich 159–179

Boyle, Matthew 239–241

Brandom, Robert 212, 214, 221

Breazeale, Daniel 90–91, 93, 101, 139

Crusius, Christian August 126, 167

Descartes, René 167, 169–170

Feder, Johann Georg Heinrich 126, 133, 167

Fichte, Johann Gottlieb 49–50, 71–74, 80,
105, 108–109, 125–136, 139–156, 162,
166, 169–170, 173–174

Forster, Georg 126

Frankfurt, Harry G. 199

Grimm, Jacob 126

Grotius, Hugo 75

Hartmann, Nicolai 136, 252

Hegel, Georg Wilhelm Friedrich 71, 73, 75,
80, 85, 109, 136, 139, 160, 203–215,
217–222, 227–239, 241–245

Heidegger, Martin 177, 181, 192

Husserl, Edmund 218

Jacobi, Friedrich Heinrich 125–127, 131–
136, 140, 160, 162–163, 166, 173, 179

Kant, Immanuel 7–26, 29–47, 49–69,
71–75, 77, 80–85, 89–99, 102, 105–
122, 130–131, 139–145, 147, 154, 156,
159, 161–172, 175–177, 179, 181–194,
196–197, 199–220, 222, 229, 235–
236, 244, 248, 258–259, 261

Kosch, Michelle 139, 154, 182, 193

Krug, Wilhelm Traugott 135

Lazzari, Alessandro 90, 92–95

Leibniz, Gottfried Wilhelm 29, 162–163,
167, 170, 173, 214

Locke, John 75, 79–80, 162, 167, 170, 173

Machiavelli, Niccolò 79, 82

Maimon, Salomon 105, 107–109, 112–122

Marx, Karl 71, 73, 178

McDowell, John 211, 238

Mendelssohn, Moses 126, 130–131

Merleau-Ponty, Maurice 218

Montesquieu, Charles-Louis de Secondat
75, 80–81

Neeb, Johann 135

Nietzsche, Friedrich 71–73, 77, 252–253

Pippin, Robert 243–244

Platner, Ernst 126, 167

Popper, Karl R. 222

Prauss, Gerold 91, 98–101

Quine, Willard V. 213

Reimarus, Hermann Samuel 126

Reinhold, Karl Leonhard 73, 89–103, 107,
127–128, 131, 140–141, 160, 162, 182,
184–187, 189–192, 196–197, 199

Rousseau, Jean-Jacques 75, 80–83, 107

Ryle, Gilbert 206–207

Schelling, Friedrich Wilhelm Joseph
49–50, 71–74, 105, 108, 174, 181–182,
189–200

Schiller, Friedrich 128

Schmid, Carl Christian Erhard 94–98,
101–102, 182

Schulze, Gottlob Ernst 162, 164

Searle, John 203

Sidgwick, Henry 184

Strawson, Galen 230–231, 236

Sulzer, Johann Georg 133

Villers, Charles François Dominique de 179

- | | | | |
|----------------------|---|----------------------|--------------------------|
| Wood, Allan | 8, 58, 62, 139, 144, 147–149,
154–155, 200, 234, 245 | Yeomans, Christopher | 232, 234–235,
244–245 |
| Xie, Simon Shengjian | 8, 12–14, 16–18 | Zöller, Günter | 71, 90–91, 99, 101, 139 |

Index of Subjects

- action 1, 8–13, 15–26, 30–41, 43,
45–46, 49, 52–53, 57–58, 60, 64, 68,
78, 83–84, 91–93, 96–100, 105–107,
110–111, 113–114, 116, 118–119, 121,
125, 128–134, 136, 140, 142–144, 148,
155, 175, 177, 182, 184–189, 191–192,
195, 197–200, 204, 206, 209, 213, 221,
229–231, 235, 240–245, 249–253, 257
- activity 30–31, 37–39, 44, 91, 98–102,
129, 134, 139, 141–156, 172, 186, 191–
192, 219, 222, 228, 233–235, 239, 242,
249, 255
- animal 61, 93, 126, 130, 132–133, 135–
136, 144, 195–196, 239
- antinomy 2, 19, 24–26, 60, 67, 149
- apperception 164, 169, 172, 174
- arbitrariness 41, 80, 219, 227, 235, 237,
239
- asceticism 247–249, 260–261
- autonomy 2, 4, 31, 41, 43–45, 50, 67–68,
72, 84, 96, 145, 147, 154, 181–186,
188–189, 192–193, 196, 199, 235–
236, 239
- body 79, 142–143, 145, 152, 193, 205,
213, 217–220, 235, 255, 260
- categories 164–165, 177, 206–209, 211
- causality 7–10, 14, 16, 19, 21, 25–26,
30–32, 36–37, 42–43, 45, 97, 130,
132–133, 136, 142–143, 147, 154, 183–
185, 196, 257
 - causality of freedom 7–9, 21, 184, 196
 - causality of nature / natural causality 9,
14, 19, 21, 25–26, 30, 42–43, 45, 130,
132–133, 136, 143
- choice 1–3, 36, 38–42, 45–47, 50,
57–59, 63–65, 69, 72–73, 79, 83–84,
89, 101, 111, 130–131, 135, 144, 155,
175, 182, 184, 186–194, 198–199, 207,
227–228, 230–231, 234, 236–237,
242, 244
- compassion 247–254, 260–261
- compatibilism 8, 13–14, 181, 194
 - incompatibilism 13
 - real compatibilism 181, 184
- consciousness 36, 39, 44–45, 53, 89–96,
98–99, 101–102, 109, 112, 114–116,
121, 128, 130, 132–133, 135–136, 142,
144, 146–148, 150–153, 156, 164, 168–
169, 171, 189–191, 194, 212, 218, 228,
252, 255
 - self-consciousness 36, 132, 135, 169,
191, 194, 218
- contingency 29, 32–35, 37, 42–43,
45–46, 121, 198, 214, 216, 233–234,
237–238
- decision 1, 60, 92–94, 99–102, 107, 182–
183, 190, 192, 197, 199–200, 230, 239
- desire 13, 23, 30, 35, 37, 44, 52–53, 55,
57, 67, 72, 84, 92–97, 100–101, 110,
113, 130, 132–135, 143, 154–155, 183,
185, 187, 191, 194–199, 204, 220, 229–
231, 233–243, 249, 252, 254–255
- determinism 7, 8, 13–14, 22, 46, 129, 186,
189, 192, 198, 228–230
- drive 1, 3, 38, 55, 73, 89, 92–93, 95, 108,
113, 115–118, 121, 125–136, 139–141,
145–156, 185, 203–205, 209–211,
213–214, 217, 219, 222, 228, 231, 233–
235, 237–239, 242–244, 249–250
 - driving-force 3, 105–113, 115, 118
 - logical drive 135, 203, 204, 205, 209,
210, 211, 213, 214, 217, 219, 222
 - selfish drive 92–93, 95, 140, 185
 - unselfish drive 73, 92–93, 95, 140
- egoism 247, 249, 252–254, 260–261
- embodiment 204
- epistemology 1, 134–135, 163, 166, 177,
213, 216, 252
- ethics 1, 3, 50, 52, 71, 73–74, 76, 82–85,
92, 96, 131, 135, 139, 141–142, 146,
156, 176–177, 214, 247–249, 252–253,
261

- evil 9–10, 15, 18, 21–22, 32, 49–51,
54–69, 106, 111–112, 136, 181–182,
184, 188–190, 192–193, 197, 200, 206,
220, 221, 236, 258–259, 261
– radical evil 49–51, 54, 58–62, 64–69,
136
- fatalism 97, 102, 182, 186, 188–190, 192,
197
- feeling 107, 128, 141, 146–151, 156, 162,
176, 178, 187, 194–196, 204–206, 211,
241
– feeling of drive 148, 151
– feeling of reality 128
– feeling of respect 107, 187, 196
– freedom 2–4, 7–14, 16–17, 19, 21,
29–34, 37–47, 49–50, 53–54, 56–57,
59, 61, 64, 71–73, 75–78, 80–85,
89–91, 93–97, 99–102, 106, 109, 111,
125, 127–129, 137, 139, 141, 149, 151–
156, 168, 171–172, 175, 178, 181–186,
188–200, 204, 206, 211, 220–221,
227–231, 234–235, 237, 245, 257
– formal freedom 154–155, 193
freedom of a turnspit 31
– freedom of the will 44, 72, 91, 96–97,
106, 175, 178, 181, 183, 185, 189, 191–
192, 194, 198–199, 228, 245
– material freedom 73, 155–156
- God 29, 32–34, 36–37, 45, 53, 63, 66,
110, 113, 170, 179
- highest good 35, 68
- idealism 2, 68, 71, 108, 125, 156, 161,
167–168, 182, 188–189, 193, 201
- imperative 9, 22–24, 34, 36, 40, 42, 45,
50, 55, 57, 61, 94, 105–106, 108, 148
– categorical imperative 22–24, 42, 45,
50, 55, 57, 61, 94, 108, 148
- imputation 17
- indifferentism 182–183, 186, 189, 198
- instinct 126, 130–131, 133, 135–136, 185,
195
- intelligible 1–2, 8–10, 14–18, 30, 32,
36–37, 58, 63–66, 89, 96–97, 102,
119, 169, 182, 186, 188–190, 192, 197,
231, 257
- knowledge 15, 41, 51–52, 54, 68, 89–90,
93, 102, 112–113, 116, 141, 146, 159,
161, 163–171, 173–176, 178, 195, 203,
227, 229, 232, 25–252, 254–255, 259,
260
- law 1–3, 7, 9–11, 16, 19, 29–34, 36–46,
50, 53–68, 71, 73–76, 79–85, 89,
91–99, 101–102, 105–107, 109–112,
116–121, 129, 155, 162–163, 182–191,
193, 196, 199–200, 207, 235–236, 257
– law of nature / natural law 16, 31,
74–75, 79, 82, 94, 183, 188,
– moral law 2–3, 32, 37–40, 45, 50,
54–66, 68, 76, 84, 89, 91–99, 101–
102, 105–107, 109–111, 116–118, 155,
183–184, 186–190, 193, 196, 199,
235–236, 268
– practical law 39–41, 55, 67
– universal law 31, 55, 67, 97
- libertarianism 7–8, 11–14, 17–19, 22, 25
- liberty 40, 71, 76, 78–84
- love 50, 55–62, 64–68, 83, 106–107,
109–110, 133, 200, 251, 253
– self-love 55–62, 64–68, 83, 106–107,
109–110, 133, 253
- metaphysics 1–4, 8–9, 32, 37–39, 53,
61, 65, 68–69, 74, 82–84, 90, 92, 113,
140, 163, 166, 177, 181, 186–187, 189,
191–192, 200, 237, 247, 252, 254, 258,
261
– metaphysics of will 181, 237, 247, 254
- methodology 227
- modality 203–205, 212, 214–215, 217
- morality 4, 17, 22, 24–26, 35, 37, 39, 47,
51–56, 60–61, 65, 67–69, 83, 97–98,
105, 107, 109–116, 118–121, 125, 134,
140, 143, 185, 191, 196, 221, 249–250,
252–253
- naturalism 135, 213
- nature 3, 7, 9, 12, 14, 17, 19, 21, 29–37,
41–43, 45, 49, 52–54, 56, 58–68, 71,

- 73, 75–76, 81–82, 97, 113–115, 117,
121, 126, 129–130, 133, 139–141, 143–
145, 148–149, 151–154, 156, 164, 166,
169, 172–174, 177, 183, 188, 191, 193–
197, 199–200, 206–207, 218, 229,
232–239, 241, 243, 245, 247–251,
253–254, 256, 259–261
- necessity 8, 19, 21, 23, 25–26, 29,
33–34, 40, 42–45, 54, 97, 161, 171,
181, 187, 190–192, 194, 198–199, 208,
214–215, 217, 220, 230, 234
- necessity of nature / natural necessity
19, 21, 23, 32, 97, 191
- volitional necessity 181, 198–199
- normativity 175, 220–221
- objectivity 118–119, 164–166, 168–170,
173, 176–177, 191, 205, 210, 213, 244
- ontology 4, 179, 203–204, 212
- social ontology 203–204, 212
- paradox 7–8, 11–12, 14, 17–19, 21,
25–26, 227, 229–232, 234–237, 243–
244, 247–249, 254, 257–259
- person 9–13, 16, 19–22, 25, 55–58, 63,
66, 79, 91, 93–94, 99–101, 120, 182,
185–186, 193–194, 196–200, 204,
220–222, 240, 251–254, 257, 261
- philosophy 1–4, 7, 11–12, 20, 25–26, 32,
47, 51, 54, 65, 69, 71–81, 83, 89–92,
96–97, 102–103, 108, 112–113, 119,
126, 139–143, 156, 159–164, 166–177,
179, 182, 185–186, 189–190, 192–194,
200, 203–204, 207–209, 211, 213–
214, 227, 229–230, 232, 235, 237–238,
241–242, 245, 247–248, 253–254,
258–261
- legal philosophy 75
- political philosophy 1, 2, 71, 74–77,
79–81, 83, 269
- practical philosophy 1–3, 11, 26, 71,
73–74, 76, 83, 89–90, 96, 139–140,
142–143, 166, 170–171, 173–176, 227,
229, 245
- Presupposition 14, 33, 44, 159, 164, 166,
177, 204–205, 211–212, 214–215, 217–
218, 220–222, 236
- constructed presupposition 205, 211–
212, 218, 220, 222
- genetic presupposition 211–212, 217,
222
- propensity 50, 53, 57–59, 61–62, 64–69,
153, 236
- psychology 1, 43, 76, 84, 89–90, 93, 135,
203, 212, 218, 220, 245
- realism 3, 161, 165, 170, 193
- reality 34, 67, 73, 93–96, 102, 116, 128–
129, 149, 159, 161, 164–176, 178, 197,
203, 212, 216, 220–222, 228, 247–
252, 254, 256, 258–261
- reason 1–2, 4, 8–12, 14, 16, 18–19,
21–26, 30–44, 46–47, 49–55, 57–61,
65–69, 71–73, 81–82, 84, 90–100,
103, 105, 107–109, 111–112, 114–121,
130, 133–135, 140, 143–145, 147, 150–
151, 159, 162–164, 166–167, 170, 172,
175, 177, 179, 181–193, 195–200, 207,
214, 219, 231, 238–240, 243–244,
248–250, 253–254, 257, 260–261
- fact of reason 38–39, 105, 108–109,
112, 114
- practical reason 1–2, 30, 38–40, 42,
44, 46–47, 54, 57, 60–61, 67, 72,
91–92, 97–100, 118–119, 133, 140,
179, 181–182, 184, 186–190, 196, 199,
238–239,
- theoretical reason 118–119
- republic 74, 79–80, 82, 119, 121
- republican / republicanism 71, 79–82,
84–85
- respect 3, 29, 38, 45, 51–52, 55, 59, 69,
83, 95, 97, 99–100, 105–107, 109–110,
113, 129, 136, 151, 153, 179, 187, 193–
196, 200, 204, 211–213, 215–216, 221,
238, 242
- responsibility 11, 57, 62, 65–66, 89, 93,
97–99, 132, 229, 231–232, 257
- right 3, 8–9, 15, 17, 22, 25, 33–35, 37,
39, 41–43, 50–51, 73–75, 79, 82, 84,
91, 96, 115, 148, 162, 184, 204, 206,

- 215, 220–221, 227, 232, 235, 237–238,
24–245, 252, 256
- salvation 247–248, 254–261
- sensibility 19–20, 40, 44, 53, 60, 67, 83,
140, 182, 185
- skepticism 60–61, 65, 114, 163, 171, 179
- spirit 33, 38, 80–81, 85, 89, 97, 107, 119,
121, 136, 155, 165, 175, 178, 191, 196,
200, 204, 207, 211–212, 214, 218–221,
227, 241
- objective spirit 204, 212, 220
- subjective spirit 204, 212, 218, 241
- spontaneity 29–33, 37, 40–41, 44, 47,
172–173, 175, 178–179, 187, 196
- striving 1, 4, 126, 135, 141, 145, 149, 151–
152, 154–155, 190–191, 247–250, 252,
254, 258, 260
- tendency 34–35, 50, 54, 56–58, 61–62,
65, 67–68, 128, 141, 145–156, 236
- thinking 36, 43–44, 71, 73–74, 76, 79,
120, 144, 147, 165, 167–169, 171–173,
176, 205, 207, 214, 230, 232, 234, 242,
245
- transcendental apperception 164, 169,
172, 174
- universality 29, 234, 237
- volition 57, 68, 84, 92–93, 100, 128, 130,
133–134, 195–196, 198–200, 228,
230, 232, 235, 244
- will 1–4, 7–14, 16, 18, 20–22, 24–26,
29–30, 32–42, 44, 46–47, 49–51, 53,
55–57, 59–69, 71–73, 75–77, 80,
82–84, 89–102, 105–113, 115–120,
125–132, 135, 139–146, 151–156, 159–
166, 171–178, 181–200, 203–206,
209, 211–214, 217–222, 227–238,
242–245, 247–261
- blind will 194
- general will 120, 195–196, 199
- objective will 191
- self-will 194–196, 198–199
- subjective will 204–206, 209, 211–213,
217, 220–222
- substantial will 204, 220, 222
- universal will 195–196, 200
- Willkür* 2–4, 30, 37, 39–40, 46–47,
49–51, 57–60, 63–64, 72, 84, 101,
106, 140–141, 154–155, 182, 186–187,
189–190, 198–199, 219, 227–239,
241–244
- Zoology / zoological 125–126, 129, 131–
132, 135

Notes on Contributors

Sorin Baiasu is Professor of Philosophy at Keele University and Director of the Keele-Oxford-St Andrews Kantian (KOSAK) Research Centre. He held visiting positions at the University of Sheffield, University of Vienna, Oxford University and the University of Warwick. He published *Kant and Sartre: Re-discovering Critical Ethics* (2011) and edited or co-edited *Politics and Metaphysics in Kant* (2011), *Kant on Practical Justification: Interpretive Essays* (2013), *Comparing Kant and Sartre* (2015), *Sincerity in Politics and International Relations* (2017) and *Kant and the Continental Tradition: Sensibility, Nature and Religion* (2019). His work was supported financially by, among others, the British Academy, the European Commission and the European research Council.

Jenny Bunker works as a Senior Lecturer in Philosophy at the University of Roehampton, having previously taught at the Universities of Oxford and Southampton. She co-edited *Thinking Through Dance: The Philosophy of Dance Performance and Practice* and her current research focuses on Schopenhauer's theory of salvation.

Alex Englander is a Wissenschaftlicher Mitarbeiter at the Chair in Epistemology, Modern, and Contemporary Philosophy at the University of Bonn. He studied theology, philosophy and the philosophy of religion at the University of Cambridge and at the University of California, Berkeley. He researches and publishes primarily on practical philosophy in the Kantian and post-Kantian traditions.

Tom Giesbers is a historian of philosophy focusing on 19th century German and 20th century French philosophy. His research is generally concerned with the intersections of theoretical and practical philosophy. Among his specializations are German idealism, German realism, German romanticism and psychoanalysis. He is an assistant professor at the Open Universiteit.

Halla Kim is professor of philosophy at Sogang University in Seoul, Korea. His areas of specialization are Kant and German Idealism as well as Korean Philosophy and Modern Jewish Thoughts. His books include *Kant and the Foundations of Morality* (Lexington Books, 2015) and he also published three anthologies, *Kant, Fichte and the Legacy of Transcendental Philosophy* (Lexington Books, 2014), *Transcendental Inquiry: Its Origin, Method, and Critiques* (Palgrave Macmillan, 2016) (both with S. Hoeltzel), and *Jewish Religious and Philosophical Ethics* (Routledge, 2017) (together with C. Hutt and B. D. Lerner).

Manja Kisner received her PhD from LMU Munich in 2016. Her dissertation on the development of the notion of will from Kant and German Idealists to Schopenhauer was published as a monograph (*Der Wille und das Ding an sich*, Königshausen & Neumann, 2016). Since 2017 she has been postdoctoral researcher and lecturer at LMU Munich. She held research stays at KU Leuven and Brown University. Selected publications: *Das Selbst und die Welt. Beiträge zu Kant und der nachkantischen Philosophie* (with G. P. Basile, A. Lyssy, M. Weiß; Königshausen & Neumann, 2019), "Fichte's Moral Psychology of Drives and Feelings and its Influence on Schopenhauer's Metaphysics of the Will" (International Yearbook of German Idealism,

2020), “Kant’s Analogy between the Moral Law and the Law of Nature” (Con-Textos Kantianos, 2019).

Markus Kohl received his PhD in philosophy from UC Berkeley in 2012. He was an assistant professor of philosophy at UT Knoxville from 2012–2017. Since 2017 he has been an assistant professor of philosophy at UNC Chapel Hill. His interests range widely over the history of philosophy and literature, with special emphasis on Kant.

Amit Kravitz (LMU, München) completed his dissertation at the Hebrew university of Jerusalem. He has published many articles on Kant and German Idealism, among others in *Kant-Studien*, *Archiv für Geschichte der Philosophie* and *Journal of the History of Ideas*. He edited the volume *Der Begriff des Judentums in der klassischen deutschen Philosophie* (Mohr-Siebeck, 2018) with Jörg Noller, and the volume *Beyond the Mind’s Limits* (Palgrave Macmillan, 2019) dedicated to Jean Améry’s work, with Yochai Ataria and Eli Pitcovski.

Ansgar Lyssy is researcher (‘wissenschaftlicher Mitarbeiter’) at the University of Heidelberg, PI with a research project on causality in Hegel, funded by the Thyssen Foundation. From 2014–18, he has been working on the concept of humanity in German Enlightenment philosophy, funded by the German Research Foundation. Before that, he was a postdoc at the Université de Montréal and a doctoral student at TU Berlin, where he received his PhD. Selected publications include: *Kausalität und Teleologie bei G.W. Leibniz* (Studia Leibnitiana Sonderheft), Stuttgart: Franz Steiner, 2016; “Kant on the vocation and formation of the human being”, in: Robert Louden / Gualtiero Lorini (eds.): *Knowledge, Morals and Practice in Kant’s Anthropology*, London: Palgrave Macmillan 2018, 81–98; and a new edition of Friedrich Bouterwek: *Idee einer Apodiktik*, Teil I (1799), with an introduction and critical notes: Stuttgart: Frommann-Holzboog, 2017.

Jörg Noller is Wissenschaftlicher Mitarbeiter and Redakteur of the “Philosophisches Jahrbuch” at the University of Munich. He completed his dissertation on the problem of individual freedom in Kant at the University of Munich, and is currently writing his Habilitation on personal life forms. He has completed research stays at the University of Notre Dame, Chicago, and Pittsburgh, and published many articles on Kant and German Idealism, among others in the *European Journal of Philosophy* and in the *Deutsche Zeitschrift für Philosophie*. His research focuses on practical philosophy in Classical German Philosophy, on the concepts of freedom and personhood, and on the problem of evil. Along with John Walsh, he is co-editor and co-translator of the volume *Kant’s Early Critics on Freedom of the Will* under contract with Cambridge University Press.

John Walsh is Wissenschaftlicher Mitarbeiter and Koordinator of the Internationale Graduiertenschule “Verbindlichkeit von Normen der Vergesellschaftung” at the Martin-Luther-Universität Halle Wittenberg. He completed his Bachelor’s Degree at the University of Kentucky, received his Ph.D. from the University of South Florida, and has completed research stays at Universität Heidelberg, Université de Fribourg and Brown University. His research focuses on practical philosophy in Classical German Philosophy. Along with Jörg Noller, he is co-editor and co-translator of the volume *Kant’s Early Critics on Freedom of the Will* under contract with Cambridge University Press.

Daniel Wenz is a research assistant at the Chair for Theory of Science and Technology at the RWTH Aachen. He studied philosophy and history at Bonn University. He graduated with a Magister thesis on the concept of nature in classical German philosophy and the contemporary theory of science. He received his Dr.phil for a thesis on speculative logic and inferential semantics from Hagen University and the Pontificia Universidad Católica Argentina (Cotutelle). He was a visiting scholar at the University of Pittsburgh (USA) and the UCA (Argentina). His research interests are classic epistemology, the theory of science and the philosophy of mathematics and logic. He is currently working on an epistemology of automated theorem proving.

Günter Zöller is Professor of Philosophy at the University of Munich. Areas: Kant, German idealism, political philosophy. Recent book publications: *The State as a Means to an End* (in German 2011), *Critical Spirit. Knowing and Acting in Kant, Fichte and Nietzsche* (in Croatia, 2012), *Reading Fichte* (in German 2013, in Japanese 2014, in Spanish 2015, in Italian 2017, in Chinese 2019), *Res Publica. Plato's "Republic" in Classical German Philosophy* (2015), *The Cambridge Companion to Fichte* (2016), *Philosophy of the 19th Century. From Kant to Nietzsche* (in German 2018), *Hegel's Philosophy. An Introduction* (in German 2020).